IDENTIFICATION OF A NOVEL PROTEIN INTERACTING WITH RPGR

JAMES BOYLAN

THESIS SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF EDINBURGH

2001

DECLARATION

I declare that

a) I have composed this thesis myself; and

b) the work is my own, except where stated.

James Boylan, B.Sc. (Hons)

ABSTRACT

X-linked retinitis pigmentosa is a severe form of human retinal degeneration and is most commonly caused by mutations in the *RPGR* gene, which maps to chromosomal region Xp21.1. The *RPGR* gene product is a novel protein containing an amino terminal domain that is homologous to the guanine nucleotide exchange factor for a small GTP-binding protein. No other domains of known function are recognisable.

The principal aim of this thesis was to identify proteins that interact with RPGR. This was done in an attempt to obtain structural or functional information about this novel protein that could account for its roles in health and disease. The yeast two-hybrid method was used to screen a variety of expression libraries made from human pancreas and bovine retina. A novel interacting protein, called RPGRIP, was identified using a bovine retina library and the human homologue identified by database screening and primer extension methods. The specificity of the interaction between RPGR and RPGRIP was confirmed by co-immunoprecipitation. Further evidence for a functionally significant interaction was obtained using six RPGR mutants, corresponding to known disease-causing mutations, all of which abolished or severely reduced the interaction.

The expression of human and bovine *RPGRIP* was examined using northern analysis and/or RT-PCR. This revealed strong expression of the human transcript in testis and retina but more widespread expression in bovine tissues. At least three alternatively spliced forms were identified. The human gene structure was identified and shown to consist of 15 exons spanning 34 kb and to code for alternatively spliced proteins of 586 amino acids, 837 amino acids, and 902 amino acids.

A structure-function analysis of RPGRIP by database searching identified two strongly predicted coiled-coil domains and two calcium-dependent phospholipid-binding C_2-domains. The gene was mapped to chromosomal region 14q11 by radiation hybrid mapping and screening a human monochromosomal hybrid panel. Mutation analysis of three autosomal recessive retinitis pigmentosa families previously shown to be linked to the 14q11 region failed to identify disease-causing mutations. The function of RPGRIP remains to be elucidated but it is a strong candidate gene for causing another form of human retinopathy.

ACKNOWLEDGEMENTS

Thank you. C.H. Sung for the bovine retina cDNA library, Julie Kenyon for a human retina, Alan Lennon for establishing lymphoblastoid cell lines, Raf Vervoort for bovine RNA samples, Judy Fletcher for help with *in situ* work, Mairi Wallace for yeast plasmids, Human Genetics Unit photography department for taking photographs of yeast colonies. Gratitude to others for advice in the lab, help with techniques, valuable comments about this thesis and so on: Sandy Edgar, Anne Seawright, Forbes Manson, Alison Brown, Alan Wright, Raf Vervoort, Brian Tulloch, Philippe Gautier, Luisa Money, Richard Axton. Thank you.

ABBREVIATIONS

3-AT	3-aminotriazole
ABC	ATP-binding cassette
AD	activation domain
ADRP	autosomal dominant Retinitis Pigmentosa
AMD	age-related macular degeneration
ARF	ADP-ribosylation factor
ARP	ADP-ribosylation factor-related protein
ARRP	autosomal recessive Retinitis Pigmentosa
BAC	bacterial artificial chromosome
BCIP	5'-bromo-4-chloro-3-indolylphosphate
BD	DNA-binding domain
bRPGR	bovine RPGR
bRPGRIP	bovine RPGRIP
bp	base pair
cAMP	cyclic adenosine monophosphate
cGMP-PDE	cyclic guanosine monophosphate phosphodiesterase
CHC	clathrin heavy chain
CORD6	autosomal dominant cone rod dystrophy
$cPLA_2$	cytoplasmic phospholipase A_2
CRalBP	cellular retinaldehyde-binding protein
Crb1	crumbs homologue 1
CRD	cone-rod dystrophy
CRX	cone-rod homeobox-containing gene
CSF	cerebrospinal fluid
CSNB	congenital stationary night blindness
C-terminus	carboxyl terminus
ddNTP	dideoxynucleotide
DEPC	diethylpyrocarbonate
del	deletion
dH_2O	distilled water
DHRD	dominant Doyne honeycomb retinal degeneration
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid

DTT	dithiothreitol
EDTA	ethylenediamine tetra-acetic acid
EFEMP1	epidermal growth factor-containing fibrillin-like extracellular matrix protein-1
ERG	electroretinogram
ESCS	enhanced S-cone syndrome
EST	expressed sequence tag
GA	Gyrate atrophy
GAP	GTPase activating protein
GCAP	guanylyl cyclase activating proteins
GEF	guanine nucleotide exchange factor
GEF	guanine nucleotide exchange factor
GGTase	geranylgeranyltransferase
GMP, GDP, GTP	guanosine mono/di/tri-phosphate
GPCR	G protein coupled receptor
GST	glutathione *S*-transferase moiety
HA	haemagglutinin
hRPGR	human RPGR
hRPGRIP	human RPGRIP
ins	insertion
IPM	interphotoreceptor matrix
IPTG	isopropylthio-β-D-galactoside
IRBP	interphotoreceptor retinoid-binding protein
kb	kilobase pairs
kDa	kiloDalton
LCA	Leber's Congenital Amaurosis
LiAc	lithium acetate
LRAT	lecithin retinol acyltransferase
LW	long wavelength
MRC HGMP RC	Medical Research Council Human Genome Mapping Project Resource Centre
mRPGR	murine RPGR
mRPGRIP	murine RPGRIP
MW	medium wavelength
MYO7A	myosin VIIA

NBT	Nitro blue tetrazolium
NMCP1	nuclear matrix constituent protein 1
NR2E3	nuclear receptor subfamily 2, group E, member 3
NRL	neural retina leucine zipper
N-terminus	amino terminus
OAT	ornithine aminotransferase
OMIM	On-line Mendelian Inheritance In Man
ORF	open reading frame
PBD	peroxisome biogenesis disorder
PCR	polymerase chain reaction
PDE-α	cyclic guanosine monophosphate phosphodiesterase alpha subunit
PDE-β	cyclic guanosine monophosphate phosphodiesterase beta subunit
PDE-δ	cyclic guanosine monophosphate phosphodiesterase delta subunit
PDE-γ	cyclic guanosine monophosphate phosphodiesterase gamma subunit
PEG	polyethylene glycol
PKC	protein kinase C
PLC	phospholipase C
PLCδ1	inositol phospholipid-specific phospholipase C delta 1 subunit
pnr	photoreceptor-specific nuclear receptor
PPRPE	preserved para-arteriolar retinal pigment epithelium
PROML1	prominin (mouse)-like-1
R^*	metarhodopsin II
RACE	rapid amplification of cDNA ends
RCC1	regulator of chromosome condensation 1
rd	retinal degeneration
rds	retinal degeneration slow
REP-1	rab escort protein-1
retGC	retinal guanylate cyclase
RGR	RPE-retinal G protein-coupled receptor
RLD	RCC1-like domain
RNA	ribonucleic acid
RNase	ribonuclease
ROM1	rod outer segment membrane protein-1

RP	Retinitis Pigmentosa
RPE	retinal pigment epithelial cell
RPGR	Retinitis Pigmentosa GTPase Regulator
RPGRIP	RPGR-Interacting Protein
RS1	X-linked retinoschisis-1
RT-PCR	reverse transcriptase-PCR
SAG	S-antigen/arrestin
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SFD	Sorsby's fundus dystrophy
STS	sequence tagged sites
SW	short wavelength
TE	Tris-EDTA
TEMED	N,N,N',N'-tetramethylethylene-diamine
TGN	trans-Golgi network
TIMP-3	tissue inhibitor of metalloproteinase-3
T_m	melting temperature
TTPA	tocopherol transfer protein alpha
TULP	tubby-like protein
Tα-GTP	active transducin
UAS	upstream activation sequence
UTR	untranslated region
X-gal	5-bromo-4-chloro-3-Indoyl-β-D-galactoside
XLRP	X-linked Retinitis Pigmentosa

CONTENTS

CHAPTER ONE

INTRODUCTION

Approximately 1 in 2000 people suffer from inherited retinal degenerations, about half of whom suffer from retinitis pigmentosa (RP) (Boughman, Conneally, and Nance, 1980; Rattner, Sun and Nathans 1999). In a study of blindness and partial sight in England and Wales between 1990 and 1991, Evans (1995) found that, as a group, the hereditary retinal dystrophies were among the most important causes of loss of vision in the 16 to 64 years (working) age-group, accounting for 11.5% of cases (Figure 1.1). This compared with glaucoma (12%), diabetes mellitus (11.9%) and degeneration of the macula and posterior pole (mostly consisting of age-related macular degeneration; 11.3%). RP is one of the most common causes of blindness in the working years of life and is inherited in autosomal dominant (AD), autosomal recessive (AR), X-linked (XL), mitochondrial and digenic modes, depending on the nature and location of the genetic lesion. XLRP tends to manifest itself in a more aggressive manner than the other genetic types, but all forms of RP commence with night blindness, followed by deterioration of peripheral vision, which can progress to legal blindness. The following sections will summarise information on the development and structure of the retina, and how this can be altered in retinal degenerative diseases.

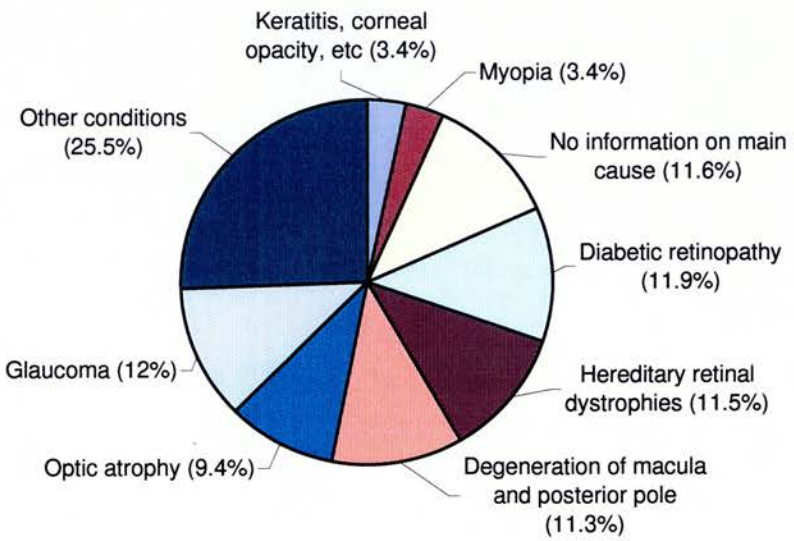


Figure 1.1: Causes of blindness in England and Wales in people aged 16 to 64 (Evans 1995).

1.1 - DEVELOPMENT AND STRUCTURE OF THE EYE

The development of the eye commences at the beginning of the fourth week of gestation. Populations of cells in the anterior neural plate of the prosencephalon (forebrain), at the cranial end of the embryo, form the so-called 'eye fields'. Optic pits, or sulci, appear in the cranial neural folds as some of the cells in the eye fields invaginate (Figure 1.2, E8.5).

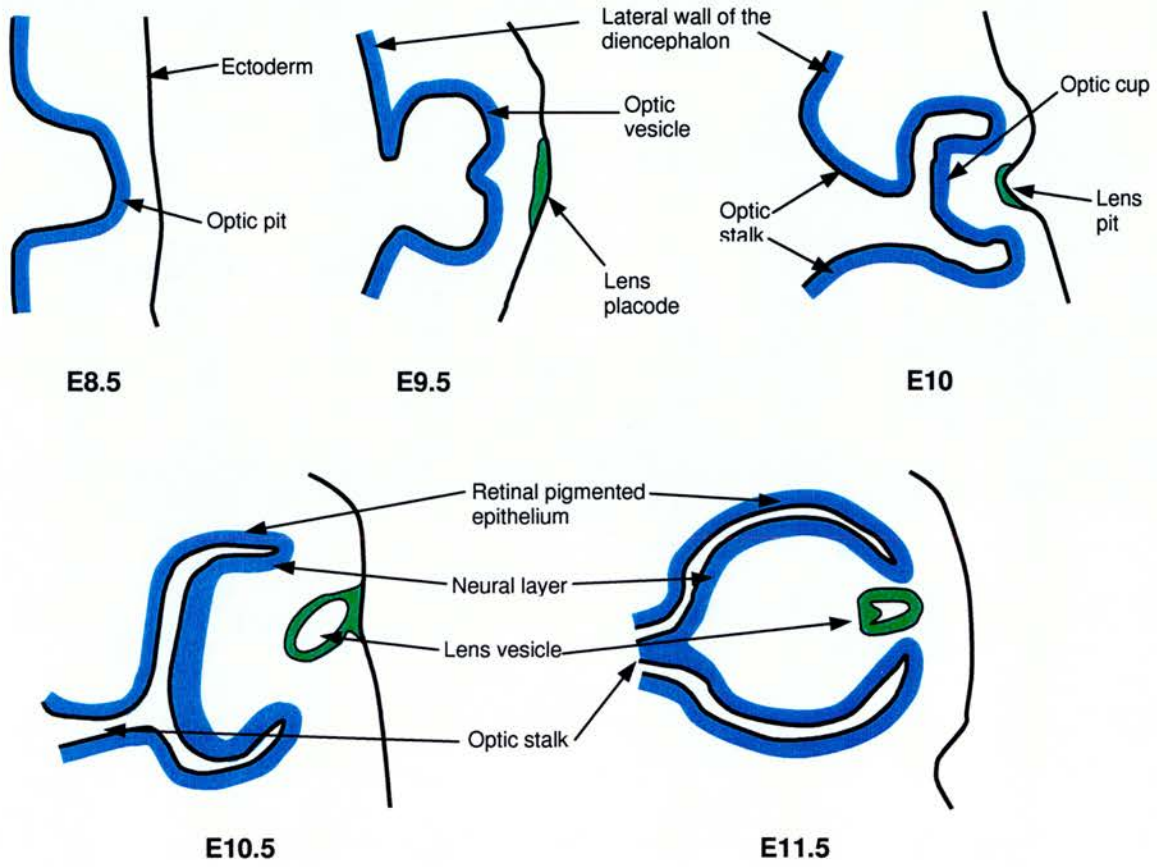


Figure 1.2: Development of the eye (numbers indicate the age of the embryo in days).

As the edges of the cranial neural folds come together to create the forebrain vesicle, the optic pits invaginate and form hollow pockets called optic vesicles. These project into the adjacent mesenchyme from the sides of the forebrain and their cavities are continuous with the lumen of the forebrain vesicle. As the optic vesicles grow they make contact with the overlying ectoderm and induce it to develop into lens tissue. The induced regions become thickened and form the lens placodes (Figure 1.2, E9.5), the central regions of which

invaginate to create the lens pits (Figure 1.2, E10). The edges of the lens pits move towards each other and fuse, so establishing the lens vesicles, which are eventually pinched off from the surface ectoderm. While this is happening, the proximal parts of the optic vesicles constrict and become the hollow optic stalks, and the lateral parts move away from the invaginating lens placodes to form the double-walled optic cups. The inner layer of the optic cup proliferates under the influence of the developing lens and develops into the neural retina (Figure 1.2, E10.5 and E11.5). The striated, laminar pattern of the retina is created by the lateral migration of postmitotic cells generated in the germinal layer. These cells differentiate into Muller glial cells and into the three major classes of neuronal cells. These are retinal ganglion cells, interneurons (bipolar, horizontal and amacrine cells) and the light-sensitive photoreceptors (the rods and cones). The outer layer of the optic cup develops into the retinal pigment epithelium. This is a non-neural layer of epithelial cells of neuroectodermal origin.

As the eye develops, the junction between the optic cup and optic stalk narrows and forms the optic disc. The ganglion cell axons in the superficial layer of the neural retina project through the optic disc and grow proximally in the wall of optic stalk. The cavity of the optic stalk becomes obliterated as a result and the ganglion cell axons form the optic nerve.

Following the development of the lens vesicle into the mature lens, the overlying surface ectoderm is induced to develop into the epithelium of the cornea and the conjunctiva. This in turn allows the immigration of mesenchymal cells of neural crest origin, which form the endothelium of the cornea. The region on the outer lip of the optic cup where the prospective neural and pigmented retinal layers meet develops into the ciliary body and the iris. The mesenchymal cells surrounding the optic cup differentiate into an outer fibrous layer, the sclera, which is continuous with the substantia propria of the cornea, and an inner vascular layer, the choroid. The vitreous body, which is composed of vitreous humor, a transparent, gel-like, avascular material, is derived from the vascular mesenchyme in the optic cup.

For recent reviews of eye development see Treisman (1999); Gilbert (1997); Morrow, Furukawa, and Cepko, (1998) and Jean, Ewan, and Gruss, (1998).

1.2 - THE RETINA

The role of the retina is to transmit neural signals conveying visual information from the eye to the brain. The RPE and neural retina together contain in excess of 100 different cell types and are divided histologically into 10 distinct layers (Figure 1.3).

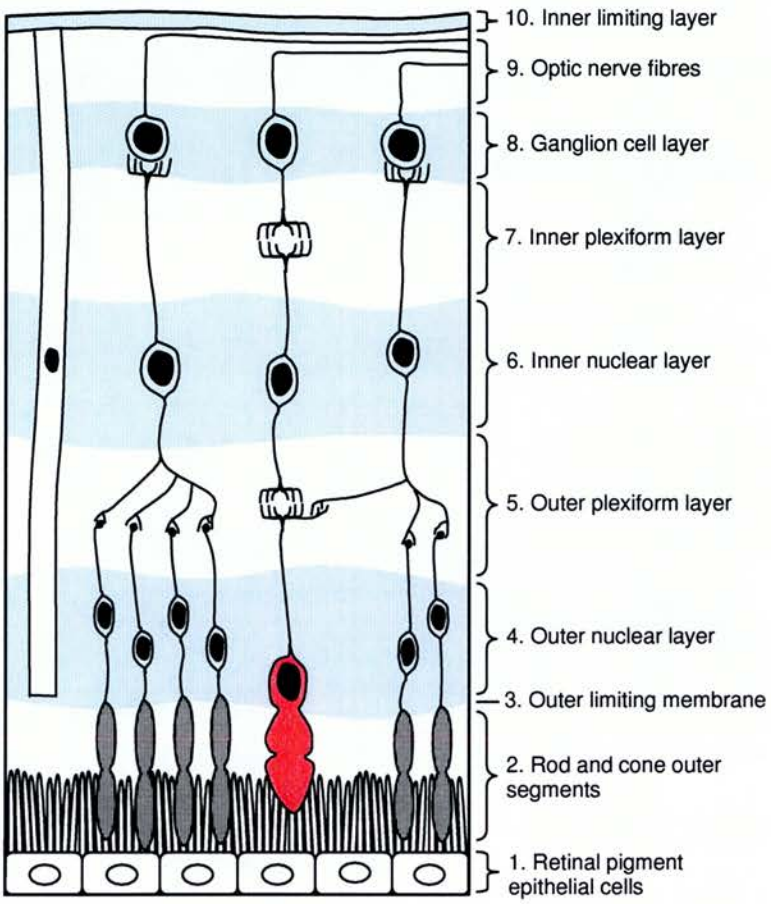


Figure 1.3: Schematic diagram showing the ten histological layers of the retina. The elongated cell on the left of the figure represents a Muller cell. Only bipolar cells are represented in the inner nuclear layer.

The outermost layer consists of retinal pigment epithelial (RPE) cells (layer 1) and is separated from the choroid by Bruch's membrane. The role of the RPE cell layer is essentially to maintain the physiology of the photoreceptors. It has several functions, including photoreceptor disc renewal, transport of water and metabolites, immunoregulation, maintenance of the blood-retinal barrier, free radical scavenging and retinal attachment. RPE cells also contain granules of melanin pigment that serve to reduce the scattering of light within the eye, so enhancing visual acuity. The choroid is the vascular coat of the eye and its blood vessels supply the outer retina via the RPE. Layer 2 consists of the outer and

inner segments of rod and cone photoreceptors. Layer 3 is the outer limiting membrane, which is not a true membrane but rather the *zonulae adherentes*, in which the Muller cell processes make contact with photoreceptor cells. The abundant Muller cells (Figure 1.3) are analogous to the neuroglia of the central nervous system and extend from the interface between the vitreous body and the junction of the inner and outer segments of rods and cones. This layer also marks the internal border of the subretinal space (the external border being the RPE cell layer). This space contains photoreceptor inner and outer segments and the interphotoreceptor matrix (IPM). The IPM is extremely narrow and contains numerous molecules important for retinal function. These include several proteoglycan species, which coat the photoreceptor outer segment, and interphotoreceptor retinoid-binding protein (IRBP), which accounts for at least 70% of the IPM protein content. IRBP is responsible for the transport of retinoids between RPE cells and photoreceptors (see section 1.5). Layer 4 is the outer nuclear layer, which contains the nuclei of rod and cone photoreceptors. Layer 5 is the outer plexiform layer, which contains rod and cone fibres and the dendrites of bipolar cells. Bipolar cells are true neurons interposed between photoreceptor cells and ganglion cells. Their dendrites contact rod photoreceptors only, or cones only, never both. Bipolar cells receive signals from varying numbers of rods (from 10, near the macula (see section 1.4), to 100 at the periphery), and while there is some convergence of cones on bipolar cells on the fringes of the human retina, this does not occur at the fovea centralis (see section 1.4). Here each cone fibre synapses with the dendrites of several bipolar cells, reflecting the higher visual acuity at this point. Layer 6 is the inner nuclear layer, which consists of the nuclei of bipolar, horizontal, amacrine, interplexiform and Muller cells. Horizontal cells are association neurons, which are capable of modifying synaptic transmission in the human retina, and are found in the outer part of the zone occupied by the cell bodies of bipolar cells. Their dendrites make contact with the synaptic terminals of photoreceptors and with the dendrites of bipolar cells. Amacrine cells are also association neurons and are located in the inner part of the zone occupied by the cell bodies of bipolar cells. They receive signals from other amacrine cells and from bipolar cells and send signals to ganglion cells, amacrine cells and bipolar cells. Interplexiform cells are a third class of association neuron, which are interspersed among the cell bodies of bipolar cells. Layer 7, the inner plexiform layer, contains the synaptic termini of inner nuclear layer neurons and dendritic processes of ganglion cells. Layer 8 is the ganglion cell layer and contains the cell bodies of ganglion cells. The axons of ganglion cells constitute layer 9, the nerve fibre layer. These axons form the optic nerve after leaving the lamina cribrosa. The ganglion cells embody the last retinal link in the visual pathway. They transmit signals to the brain in the form of a continuous

background rhythm of action potentials. The visual signal is superimposed on this background, either increasing or decreasing the rate of nerve impulses. Layer 10 is the inner limiting membrane, which is formed by the expanded inner processes of Muller cells and separates the retina from the vitreous body.

1.3 - ROD AND CONE PHOTORECEPTOR CELLS

Photoreceptors are the primary neurons of the visual pathway. Photoreactive pigments within these cells initiate the transduction of energy stored in photons into electrical signals. There are two types of photoreceptors, called rods and cones (Figure 1.4). Rods are sensitive to

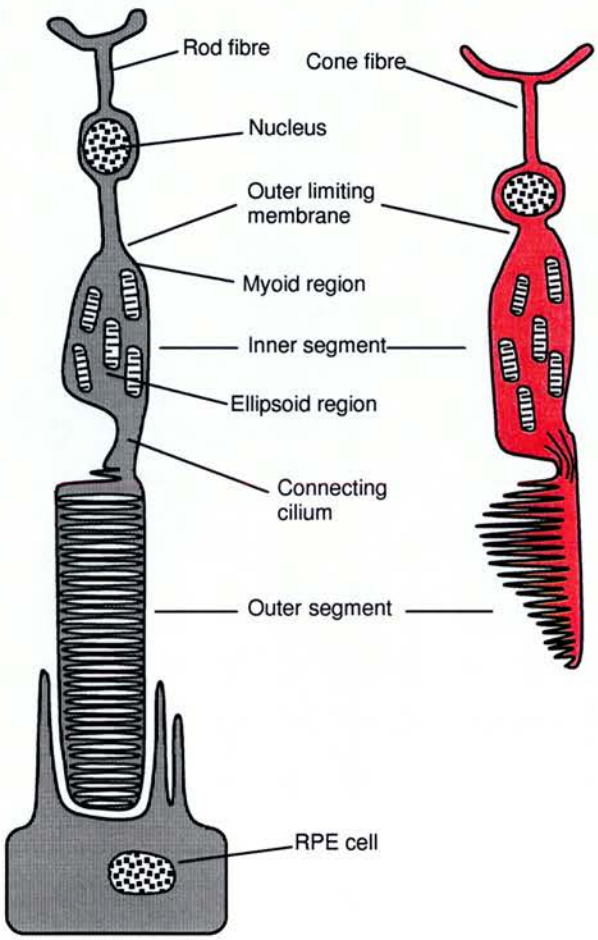


Figure 1.4: Schematic diagram of rod (grey, left) and cone (red, right) photoreceptors.

dim light, to movement and shapes, whereas cones mediate colour vision. The central portion of the retina, the macula, is the area providing highest acuity of vision and is the

principal part of the retina used for colour discrimination. The centre of the macula is called the fovea centralis, an area only containing cones (Figure 1.5).

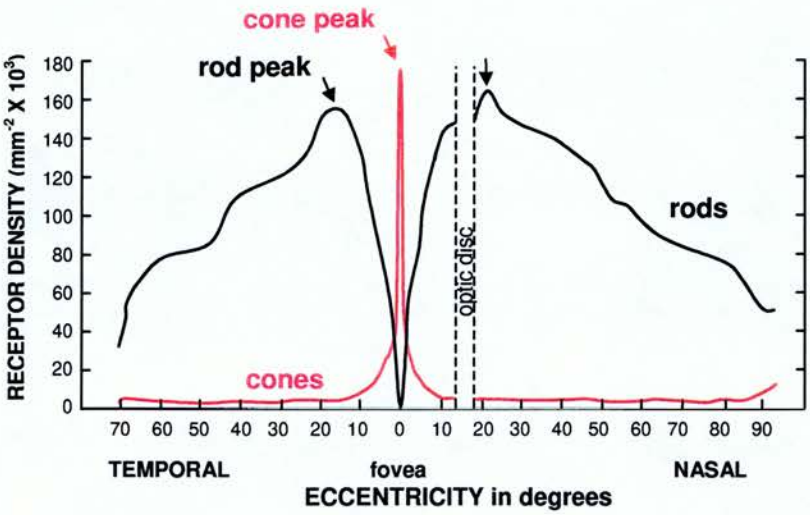


Figure 1.5: Distribution of rod and cone photoreceptors in the human retina (Osterberg, 1935).

The human retina contains approximately 130 million rods, almost 20 times the number of cones. The retina remains rod-free until approximately 0.28 mm from the centre of the fovea centralis and rod density increases from here to the ora serrata, an irregular border representing the functional anterior perimeter of the retina (Figure 1.5). Because of their more peripheral distribution (compared to cones) rods are important for peripheral vision. Rods are maximally sensitive in dim light, while cones mediate vision in ambient light, centrally and peripherally. Rods are slender cells, consisting of three portions: the outer segment, the inner segment and the nucleus and synaptic terminus (rod fibre). The outer and inner segments have a combined length of 40 μm at the periphery and 60 μm near the fovea centralis and are approximately 2 μm thick. The rod fibre is a slender thread that contains the nucleus in a central expansion and terminates as an end bulb that makes synaptic contact with bipolar and association neurons. The light sensitive part of the rod is the outer segment, a modified cilium extending from the apical surface, which contains the phototransduction proteins including the pigment rhodopsin. Most of the outer segment is occupied by 700-1000 closely spaced, double layered membranous discs, approximately 2 μm in diameter and 14 nm thick. These derive from the cell membrane but, unlike in cones, they are independent of the cell membrane and of each other. The discs are formed at the base of the outer segment and shed at the tip in a circadian manner. Pseudopodia extend from the underlying

RPE cells so that each RPE cell is in contact with approximately 45 photoreceptor outer segments. Shed discs are phagocytosed by the RPE cells and quickly degraded. Shedding of the discs occurs continuously and, in primates, the entire outer segment is renewed every 10 days or so. The inner segment contains the organelles of the photoreceptor. The ellipsoid is a region adjacent to the outer segment and contains many elongated mitochondria. From one of two centrioles at the junction of the inner and outer segments, a cilium 0.2-0.3 μm in diameter, consisting of 9 pairs of microtubules, extends into the outer segments. The central pair of single microtubules, dynein spokes and radial arms characteristic of motile cilia are absent here. The cilium is encircled by 9 to 12 microvilli extending from the plasma membrane of the inner segment. These surround the base of the outer segment and constitute the calyx. The remainder of the inner segment, the myoid region, contains vesicles, neurofibrils and granular endoplasmic reticulum. The myoid is so called because in lower vertebrates it is able to contract in response to changes in light intensity.

Cone photoreceptors are less abundant than rods (approximately 7 million in the human retina) but are important for visual acuity and colour vision. There are an estimated 100,000 cones in the fovea centralis (with 35,000 of these in the central area, the foveola) and their density decreases toward the periphery of the retina. They have a longer and thinner shape at the fovea centralis than elsewhere: 75 μm long and 1 μm thick centrally compared to 40 μm long and 6 μm thick at the periphery. Cones are also made up of an outer segment, inner segment and cone fibre. Cone fibres are thicker than rod fibres and contain the nucleus in an expanded portion next to the inner segment. They terminate with a bulbous end that makes synaptic contact with bipolar and association neurons. In addition, processes from cone synaptic termini extend to form interphotoreceptor junctions with other cones and (via gap junctions) with rods. The light-sensitive outer segment is tapered and contains a varying number of double layered membranous discs, from 1000 in foveal cones to several hundred at the periphery. Cone discs are not produced in the same way as rod discs and the rhythm of the circadian phagocytosis by RPE cells is different from rods (cones are shed at night, following light offset whereas rods are shed after light onset). They are stacked more closely than in rods, and unlike rod discs, which are surrounded by a plasma membrane, they are in free communication with the interphotoreceptor space. There are three types of pigment found in cone outer segment discs, commonly referred to as blue (short wavelength, SW), green (medium wavelength, MW) and red (long wavelength, LW). These are each maximally sensitive to a specific light wavelength: 430-440 nm for the blue pigment, 535-540 nm for green and 560-565 nm for red. Cone inner segments are thicker than the equivalent structures in rods and contain a larger ellipsoid and a more dense

granular endoplasmic reticulum. The cone cilium is comparable to that of rods. For more comprehensive descriptions of photoreceptors see Ahnelt (1998), Forrester *et al.* (1995) and Molday (1998).

1.4 - PHOTOTRANSDUCTION

The conversion of light energy into electrical signals is achieved by a series of amplification steps within the phototransduction cascade (Figure 1.6). Central to this is the visual pigment, consisting of an integral membrane apoprotein bound to a chromophore, 11-*cis* retinal. The cascade is initiated in rods when the absorption of a single photon by the 11-*cis* retinal chromophore of rhodopsin causes it to isomerise to the all-*trans* form. This leads to a conformational change in the rhodopsin protein to a form called metarhodopsin II (R^*), which, in turn, catalyses the exchange of GDP for GTP on the heterotrimeric G protein transducin. The resultant dissociation of the catalytic α-subunit from the β- and γ-subunits

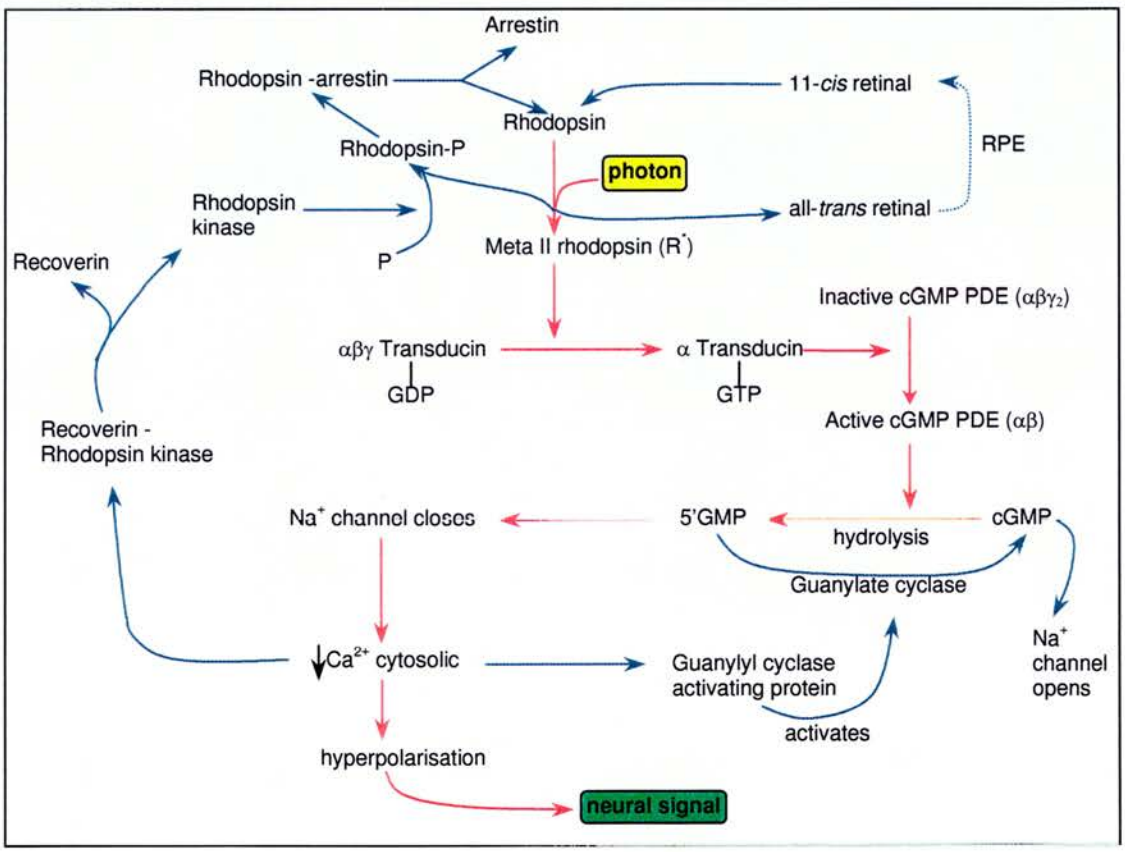


Figure 1.6: Photoexcitation and recovery of components of the phototransduction cascade. Red arrows indicate phototransduction events, blue arrows indicate photorecovery events.

causes transducin to enter the active state (Tα-GTP). The initial signal is amplified at this stage as a single meta II rhodopsin molecule activates several hundred transducin molecules. Tα-GTP interacts with cyclic guanosine monophosphate phosphodiesterase (cGMP PDE) to release the inhibitory γ-subunits from the $\alpha\beta\gamma_2$ holoenzyme. Active cGMP PDE in turn catalyses the hydrolysis of cGMP to 5'GMP, and the signal is again amplified. The result of this cascade is to reduce the level of cGMP in the cell, which in turn closes the cGMP-gated cation channels in the photoreceptor plasma membrane. In the dark, the cGMP-gated channels transport sodium and calcium ions into the outer segment (the "dark current") while other channels continuously export sodium and calcium. This leads to a partial depolarisation of the photoreceptor and, consequently, a steady neurotransmitter release. Closure of the cGMP-gated channels leads to a reduction in intracellular calcium and sodium and the photoreceptor cell membrane becomes hyperpolarised. Photoreceptor hyperpolarisation slows the ongoing release of glutamate at the synaptic terminal and so generates a neural signal (see reviews by Molday (1998), Pepe (1999), and Rispoli (1998)).

1.5 - PHOTORECOVERY AND THE VISUAL CYCLE

Calcium-binding proteins including recoverin and guanylyl cyclase activating proteins (GCAP I, II) mediate the recovery phase by responding to falling intracellular calcium $[Ca^{2+}]_i$ following a bleach of light. Recoverin inhibits the activity of rhodopsin kinase when it is in its calcium-bound state. After a bleach of light the decline in $[Ca^{2+}]_i$ releases calcium from recoverin and the inhibition of rhodopsin kinase is relieved. Rhodopsin is deactivated when it is phosphorylated at its C terminus in an ATP-dependent manner by rhodopsin kinase and becomes bound to the protein arrestin. Arrestin-binding prevents further transducin activation and permits the release of 11-*trans* retinal from rhodopsin. This rapid shut down of the cascade mechanism prevents further cGMP PDE stimulation. GCAP, which is inactive in the calcium-bound form, becomes active as $[Ca^{2+}]_i$ drops and stimulates a membrane-bound retinal guanylate cyclase to increase synthesis of cGMP. The cGMP-gated channels can then bind cGMP and open once more, restoring the depolarised state. In addition to these calcium-responsive elements, the activation of PDE is abolished following the hydrolysis of GTP to GDP on Tα, catalysed by its intrinsic GTPase activity. The transducin α-subunit reassociates with the βγ-subunits to form the inactive heterotrimer (see reviews by Jindeová (1998) and Scott and Zuker (1997).

One of the major functions of the RPE is to regenerate 11-*cis* retinal from all-*trans* retinal, which is released from the photoreceptor discs for this purpose. All-*trans* retinal

appears to be transported out of the discs following light exposure as a complex with phosphatidylethanolamine (Weng *et al.*, 1999) by the ABCA4 (ABCR) protein and then reduced to all-*trans* retinol by all-*trans* retinal dehydrogenase (Rattner, Smallwood and Nathans, 2000). Retinoid-binding proteins associate with the hydrophobic retinoids in order to solubilise them and to protect them from oxidation. Transport of all-*trans*-retinol from the rod outer segments to the RPE cells is mediated by IRBP. Once it has entered the RPE, the all-*trans*-retinol is esterified by lecithin retinol acyltransferase (LRAT) to form all-*trans*-retinyl esters and then converted to 11-*cis* retinol by an RPE-specific isomero-hydrolase, following which it becomes bound to the cellular retinaldehyde-binding protein (CRalBP). The RPE65 protein appears to be a component of the isomero-hydrolase step, although its precise role is unknown (see section 1.6.3.i). 11-*cis* retinol is then oxidised to 11-*cis* retinal by 11-*cis* retinol dehydrogenase in association with CRalBP and is subsequently released from the RPE cells and transported to the outer segments, at which stage it is again associated with IRBP, where it regenerates rhodopsin.

1.6 - GENE DEFECTS IN RETINAL DEGENERATIVE DISEASES

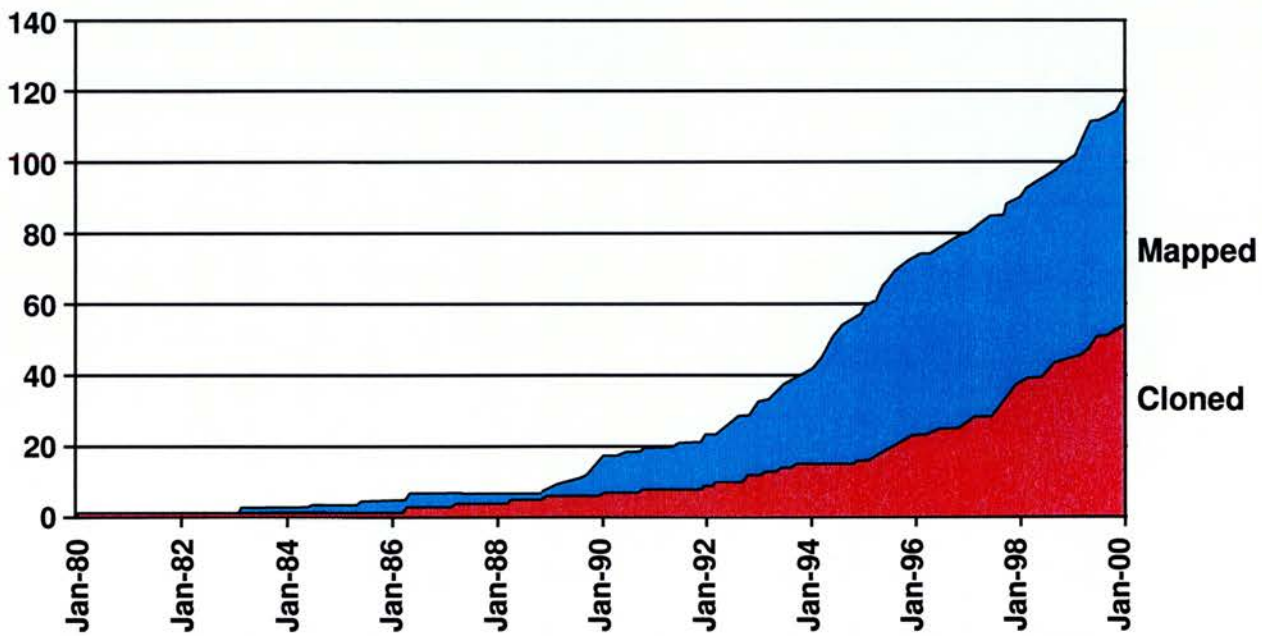


Figure 1.7: Retinal disease genes that were mapped and cloned between 1980 and 2000 (from RetNet, the Retinal Information Network http://www.sph.uth.tmc.edu/RetNet/).

In April 2000, the number of genes containing one or more allelic variants associated with a human disease recorded in the On-line Mendelian Inheritance In Man (OMIM) database

reached 1,000 (Antonarakis and McKusick, 2000). This is an up-to-date catalogue of human genes and genetic disorders (http://www.ncbi.nlm.nih.gov/omim). 98 different OMIM gene entries contain allelic variants that have been identified in association with ocular phenotypes. The number of retinal disease genes that have been mapped over the last twenty years is in excess of 120 (Figure 1.7). A large number of these genes encode either structural components of the photoreceptors and other parts of the retina, or the biochemical components of vision such as the proteins involved in phototransduction, photorecovery and the visual cycle. In the following sections, the genes that are known to be mutated in monogenic retinal degenerative diseases are described. The functions of many of the gene products are largely unknown, although hints tend to be provided by homology to characterised proteins or the presence of motifs found in other proteins. Most of the genetic disorders listed here are untreatable (with the exceptions of Refsum disease and abetalipoproteinaemia). The genes are presented in six groups according to their proposed functions: phototransduction, outer segment structure, the visual cycle, transcription factors, metabolism and miscellaneous.

1.6.1 - PHOTOTRANSDUCTION

1.6.1.i -Rhodopsin

Rhodopsin is the visual pigment found in rod photoreceptors. It functions in dim light and absorbs light maximally at a wavelength of 495 nm (Nathans *et al.*, 1984). The gene encoding human rhodopsin was isolated and characterised by Nathans and Hogness (1984) who showed that it maps to 3q21-q24 and that the amino acid sequence is 41% identical to those of the three colour photopigments (Nathans *et al.*, 1986a). Rhodopsin is a member of the G protein coupled receptor (GPCR) superfamily and contains seven transmembrane α-helices, oriented perpendicular to the plane of the disc membrane, that constitute up to 60% of its secondary structure (Pepe, 1999). The crystal structure of rhodopsin was determined recently by Palczewski *et al.* (2000). They showed that the seven-helix transmembrane region of the protein has its basis in the highly organised structure of the extracellular region and that the 11-*cis* retinal chromophore maintains this transmembrane region in an inactive conformation. Mutations in the rhodopsin gene cause dominant retinitis pigmentosa (RP), recessive RP and dominant congenital stationary night blindness (CSNB). More than 100 different disease-causing rhodopsin mutations are known (Dryja and Li, 1995) and those causing autosomal dominant RP account for 25% of cases (Dryja *et al.*, 2000). The majority

of rhodopsin mutations fall into one of two categories: class I or class II. Class I mutant rhodopsin proteins are indistinguishable from wild-type rhodopsin *in vitro* in all characteristics examined. However, in transgenic mice carrying the Q344ter class I mutation, rhodopsin is inefficiently localised to the outer segment while the endogenous wild-type rhodopsin is processed normally (Sung *et al.*, 1994). Another class I mutation, P347S, causes rhodopsin to accumulate in vesicles at the outer segment base (Li *et al.*, 1996). Class I mutations are found at the carboxyl (C) terminus, which is involved in protein sorting and transport to the outer segment. The C-terminus of rhodopsin binds to the protein Tctex-1, a widely expressed dynein light chain abundant in photoreceptor inner segments. Several autosomal dominant RP-causing class I mutations abolish binding to Tctex-1 and it is thought that this prevents the proper guidance of post-Golgi vesicles carrying rhodopsin to the outer segment (Tai *et al.*, 1999).

Class II mutant rhodopsins are defective in terms of thermal stability and/or protein folding, so that they fail to fully regenerate with 11-*cis* retinal and are retained in the endoplasmic reticulum (Liu, Garriga and Khorana, 1996; Min *et al.*, 1993; Sung *et al.*, 1993). These types of mutations usually affect the transmembrane and intradiscal domains of rhodopsin. In transgenic mice bearing the T17M class II mutation, the administration of vitamin A supplements led to a decrease in the rate of retinal degeneration (Li *et al.*, 1998). The same treatment had no significant effect on mice carrying the P347S class I mutation. This work extended the findings of Berson *et al.* (1993) showing that vitamin A supplements retard the progress of retinal degeneration in adults with retinitis pigmentosa. In binding to the 11-*cis* retinal chromophore, the thermal stability of wild-type rhodopsin is increased (Hubbard 1958). It is thought that the increased availability of vitamin A stabilises the T17M mutant rhodopsin. Evidence supporting the latter proposal came from the observation that the stability of the T17M mutant opsin expressed *in vitro* was increased by including 11-*cis* retinal in the culture medium (Li *et al.*, 1998).

1.6.1.ii - cGMP Phosphodiesterase

Rod cGMP (type 6) PDE is a heterotetrameric protein consisting of α, β and two γ subunits. The α subunit of transducin, when bound to GTP, induces the gamma subunits to detach from the inactive $\alpha_1\beta_1\gamma_2$ heterotetramer, leaving the enzymatically active α-β dimer.

Mutations in the gene encoding the 90 kDa PDE-β protein, located in 4p16.3 (Altherr *et al.*, 1992), are found in 3-4% of recessive RP cases (McLaughlin *et al.*, 1995). Mutations in the N-terminal half of PDE-β are responsible for dominant CSNB in a single

family (Gal *et al.*, 1994) whereas the majority of the autosomal recessive RP-associated mutations are found in the C-terminal catalytic domain (McLaughlin *et al.*, 1995). The retinal degeneration (rd) mouse carries a mutated *PDE*-β gene and, in homozygotes, the rod photoreceptor cells start to degenerate at postnatal day 8 and disappear by 4 weeks of age (Bowes *et al.*, 1990). Subretinal injection of an adenovirus construct containing wild-type murine PDE-β cDNA 4 days postnatally led to the production of detectable PDE-β transcripts and enzyme, and a short-lived retardation of photoreceptor degeneration (Bennett *et al.*, 1996).

The 90 kDa alpha subunit is encoded by a 22 exon gene in 5q31.2-q34 (Pittler *et al.*, 1990). Because null mutations in the *PDE*-β gene were known to cause retinitis pigmentosa and since both α and β subunits of the protein are required for catalytic activity, Huang *et al.* (1995) searched for and found mutations in the *PDE*-α gene in RP patients. Such mutations have been estimated to cause 3-4% of cases of recessive RP in North America (Dryja *et al.*, 1999).

No disease-causing mutations have been found in the *PDE*-γ gene, which is located in 17q21.1. However, disruption of the murine *PDE*-γ gene resulted in a rapid retinal degeneration similar to retinitis pigmentosa (Tsang *et al.*, 1996).

The PDEδ subunit is a highly conserved protein (Hurwitz *et al.*, 1985), the precise role of which in photoreceptors is not understood. It is known to solubilise the membrane-bound PDE holoenzyme (Florio, Prusti, and Beavo, 1996) and the addition of a recombinant epitope-tagged form of the protein to a preparation of permeablised bovine rod outer segments was recently shown to reduce the maximal rate of light-induced cGMP hydrolysis (Cook *et al.*, 2000). It maps to chromsomal region 2q35-q36 and so far no mutations in the *PDE*δ gene have been found in patients with retinal degenerative disorders.

1.6.1.iii - cGMP-gated channel-α

The rod cGMP-gated channel is a heterotetramer composed of homologous alpha and beta subunits. These each have a pore region, 6 putative transmembrane domains and cytoplasmic N- and C-termini. The gene encoding the alpha subunit is located at 4p12-cen and mutations affecting it are responsible for an estimated 1-2% of autosomal recessive RP (Dryja *et al.*, 1995). Known mutations include nonsense mutations situated near the start of the open reading frame, a deletion which removes almost the whole of the open reading frame, a frame-shift mutation and missense mutations, suggesting that RP results from loss-of-function. Frame-shift and missense mutations were investigated *in vitro* and found to

prevent the proteins from being correctly targeted to the plasma membrane (Dryja *et al.*, 1995). Mutations in the cone cGMP gated channel alpha subunit gene (*CNGA3*) at chromosomal region 2q11 have been found in patients with autosomal recessive rod monochromacy, a non-progressive disease marked by an absence of cone function (Kohl *et al.*, 1998).

The cone *CNGB3* gene, encoding the cone cGMP-gated channel β3 subunit, is localised in chromosomal region 8q21-q22 and is mutated in another form of recessive rod monochromacy (achromatopsia). This disorder is found in 4 to 10% of the Pingelapese people of the Eastern Caroline Islands and the symptoms include nystagmus, complete colour blindness and photophobia (Sundin *et al.*, 2000).

1.6.1.iv - Retinal Guanylate Cyclase and Guanylate Cyclase Activating Proteins

The retinal guanylate cyclases (retGC-1 and retGC-2) are the enzymes responsible for the formation of cyclic-GMP. In the photoreceptor cell, the level of this second messenger is regulated both by guanylate cyclase and by cGMP-PDE hydrolysis. Guanylate cyclases are either of the particulate (membrane) type or the soluble type. Membrane guanylate cyclases, of which the retinal enzymes are examples, are located in outer segment disc membranes and consist of a cytoplasmic protein kinase homology domain, a C-terminal catalytic domain and a ligand-binding N-terminal portion. Mutations in the gene encoding retGC-1 (designated *GUCY2D*), which maps to 17p13.1, cause recessive Leber's congenital amaurosis (LCA) (Perrault *et al.*, 1996) and autosomal dominant cone-rod dystrophy (Kelsell *et al.*, 1998). LCA is a group of autosomal recessive retinopathies and is the most common cause of severe inherited visual impairment in infants and children. The symptoms of LCA are seen within a few months of birth and include severe visual impairment, nystagmus, and an absent or poorly recordable electroretinogram (Perrault *et al.*, 1999). Besides *GUCY2D*, mutations at additional loci are responsible for LCA (see sections 1.6.3.i (*RPE65*), 1.6.4.i (*CRX*) and table 1.1 (*AIPL1*)). It is not known why a defect in retGC-1 is not compensated by the presence of a functional retGC-2. Interestingly, two missense mutations in codons 837 and 838 of the *GUCY2D* gene (encoding residues within the retGC-1 dimerisation domain) were found to cause autosomal dominant cone rod dystrophy (CORD6). Heterozygous carriers of recessive null *GUCY2D* mutations do not display a phenotype, suggesting that the CORD6 mutations impart a gain-of-function and that the cone-rod dystrophy is not due to haploinsufficiency (Kelsell *et al.*, 1998). The biochemical characteristics of the R838C *GUCY2D*/CORD6 mutation were investigated *in vitro* by Tucker *et al.* (1999). They found that the mutation

increases the affinity for GCAP-1 and modifies the Ca^{2+}-sensitivity of the GCAP-1 response such that mutant retGC-1 is activated by GCAP-1 at higher concentrations of Ca^{2+} than wild-type. Wilkie *et al.* (2000) investigated the effects of additional *GUCY2D*/CORD6 mutations at codon 838 *in vitro* and found a direct correlation between disease severity and residual GCAP-1-stimulated activity at high Ca^{2+} concentrations.

The guanylate cyclase activating proteins, GCAP-1 and GCAP-2, stimulate the synthesis of cyclic GMP by retGC in photoreceptor cells. The mammalian GCAP proteins, which are greater than 90% similar, consist of 201-205 amino acids, and contain three calcium binding sites (Subbaraya *et al.*, 1994). They share homology with recoverin, another photoreceptor calcium-binding protein. The two GCAP genes in humans, *GUCA1A* and *GUCA1B*, are arranged in a tail-to-tail array in chromosomal region 6p21.1 (Surguchov *et al.*, 1997). Twenty-seven members of a four generation family with autosomal dominant cone dystrophy were found to carry a Tyr99Cys (Y99C) mutation in the *GUCA1A* gene (encoding GCAP-1) (Payne *et al.*, 1998). The fact that *GUCA1A* mutations cause degeneration of cones may be because of the higher relative expression of this gene in cones than rods (Howes *et al.*, 1998).

1.6.1.v - Arrestin

Arrestin (also known as S-antigen (SAG) and 48-kDa protein) is a 405 amino acid protein involved in the recovery phase of phototransduction. Following photoactivation, arrestin works together with rhodopsin kinase to shut down rhodopsin activity: rhodopsin kinase phosphorylates rhodopsin, which is then bound by arrestin. The *SAG* gene shows homology to bovine alpha transducin and is located in chromosomal region 2q37 (Calabrese *et al.*, 1994). Mutations in this gene cause Oguchi disease and RP. Oguchi disease is a rare, autosomal recessive form of CSNB. In this disorder, light adaptation occurs very slowly in rods, while the process in cones occurs as normal. One characteristic of Oguchi disease is the Mizuo-Nakamura phenomenon, whereby the fundus undergoes a light-induced metallic golden-yellow colour transformation. The *SAG* gene was first implicated in Oguchi disease when Fuchs *et al.* (1995) identified a homozygous frame-shift mutation at codon 309 in Japanese patients. It is interesting to note that this same mutation is also found in a small percentage of Japanese ARRP patients (Nakazawa *et al.*, 1998). In one example, described by Nakazawa *et al.*, one of two siblings carrying the frame-shift mutation at codon 309 had Oguchi disease and the other had classic RP without the Mizuo-Nakamura phenomenon.

1.6.1.vi - Rhodopsin Kinase

Rhodopsin kinase is a highly specific protein kinase that initiates the recovery phase of phototransduction by catalysing the phosphorylation of rhodopsin. The gene encoding rhodopsin kinase maps to 13q34 (Khani *et al.*, 1996). Following the identification of mutations in the *SAG* gene in Oguchi disease patients, it was postulated that, since rhodopsin kinase has a similar function, mutations in this gene may also lead to Oguchi disease. This was found to be the case by Yamamoto *et al.* (1997).

1.6.1.vii - Transducin

Transducin is a heterotrimeric G-protein that transduces the light-response signal from photoexcited rhodopsin to cGMP PDE. A G38D mutation in the alpha subunit of transducin has been found in autosomal dominant CSNB patients within a single French family (Dryja *et al.*, 1996). Codon 38 is a highly conserved residue within a region that is important for GTP hydrolysis (Lambright *et al.*, 1994). In addition to retinal cells, rod transducin is also found in cells concerned with taste, and is related to Gustducin, which is a G-protein specifically found in taste receptor cells (Ruiz-Avila *et al.*, 1995).

1.6.1.viii - Cone Pigments

The cone visual pigments also consist of a cone opsin apoprotein covalently linked to the 11-*cis* retinal chromophore. Three different cone pigments exist, each with a different absorption spectrum. A single cone will contain only one pigment type, so there are three types of cone cell: red (LW), green (MW) and blue (SW). The differences in absorption spectra are thought to be attributable to amino acid sequence differences in the pigment apoproteins. At the protein level, the cone pigments are 41% identical to rhodopsin, with the red and green pigments sharing 96% identity with each other and 43% with the blue pigment (Nathans *et al.*, 1986a). The blue pigment gene is autosomal and the red and green pigment genes are located on the X chromosome. In approximately 8-10% of Caucasian males, rearrangements affecting the red and green pigment gene array produce stationary defects in colour vision (Nathans *et al.*, 1986b). Less frequently, loss of both red and green pigment function can produce blue cone monochromacy or a variant of this (Nathans *et al.*, 1989; Nathans *et al.*, 1993; Reichel *et al.*, 1989), disorders that are characterised by a loss of visual acuity and severe colour blindness. In a number of these patients there is a progressive

deterioration of the central retina, which is potentially analogous to the peripheral degeneration seen in RP patients with mutations in the rhodopsin gene (Nathans *et al.*, 1993). Autosomal dominant blue colour-blindness (tritanopia) is a stationary disorder caused by mutations in the transmembrane domains of the blue pigment gene in chromosomal region 7q31.3-q32 (Weitz, Went and Nathans, 1992).

1.6.2 - OUTER SEGMENT STRUCTURE

1.6.2.i - RDS and Rod Outer Segment Membrane Protein-1 (ROM1)

RDS (formerly peripherin/RDS) and ROM1 are homologous proteins which interact to form oligomeric complexes. They are thought to be involved in the maintenance of the membrane curvature at the rod outer segment disc rim (Goldberg and Molday, 1996; Molday, 1998). The gene encoding RDS maps to 6p21.2-cen (Travis *et al.*, 1991) and its orthologue is mutated in the *retinal degeneration slow* (*rds*) mouse, which displays a degenerative retinopathy (Travis *et al.*, 1989). Over 40 *RDS* mutations have been found in patients suffering from a variety of dominantly inherited retinal diseases, including dominant RP, dominant macular degeneration, RPE pattern dystrophies and dominant adult vitelliform macular degeneration (Keen and Inglehearn., 1996). A simple genotype-phenotype correlation for different *RDS* alleles does not exist. For example, different members of the same family with a single *RDS* mutation suffer from RP, pattern dystrophy and fundus flavimaculatus (Weleber *et al.*, 1993). Also, in several families, individuals carrying a L185P mutation in the *RDS* gene and a null *ROM1* allele suffer from RP while individuals carrying either mutation alone are not affected (Dryja *et al.*, 1997). This was the first example of a digenic disorder in man.

The *ROM1* gene maps to 11q13 and became an obvious candidate gene for monogenic retinopathies (Bascom *et al*,. 1992). Although *ROM1* sequence variants have been found among RP patients, a causative relationship with the disease has not been established (Bascom *et al.*, 1993). However, in *Rom1*(-/-) mice, the rod outer segments were found to be highly disorganised, with unusually large discs, and the rod photoreceptors slowly apoptosed. This suggested that the rodent *Rom1* gene is essential for the viability of rod photoreceptors and for the regulation of outer segment disc morphogenesis.

1.6.2.ii - Myosin VIIA

Usher syndrome affects approximately 1 in 25,000 people and is a genetically heterogeneous autosomal recessive group of disorders characterised by congenital hearing loss and RP (Boughman *et al.*, 1983). Three subtypes are separated according to severity of hearing loss and vestibular dysfunction: type I patients lack vestibular function and suffer a profound deafness, type II patients have normal vestibular function with a less severe hearing loss and type III patients have an inconsistent vestibular function and progressive hearing deficiency. In developed countries, Usher syndrome is the most common cause of combined deafness and blindness (Boughman, Vernon, Shaver, 1983). Several loci are involved in type I Usher syndrome but type 1B, which maps to 11q13.5, accounts for approximately 75% of type 1 patients (Weil *et al.*, 1995). The *myosin VIIA* (*MYO7A*) gene is mutated in Usher syndrome type 1B (Weil *et al.*, 1995). Myosin VIIA is an unconventional myosin with a predicted molecular weight of approximately 250 kDa. It was localised in primates to the apical microvilli of the RPE and to the ciliary base of rod and cone outer segments (Liu *et al.*, 1997). A role has been proposed for myosin VIIA in outer segment biogenesis (El-Amraoui *et al.*, 1996; Liu *et al.*, 1997).

1.6.2.iii - REP1

Choroideremia is an X-linked disorder, similar to RP, which affects some 1 in 50,000 males (Cremers and Rogers 1995). At an early age, patients experience night blindness which progresses in middle life to visual field loss and choroido-retinal atrophy (Heckenlively and Bird, 1988). The gene encoding subunit A of the rab geranylgeranyltransferase (GGTase) enzyme was found to be mutated in choroideremia by positional cloning (Cremers *et al.*, 1990). This protein, which is also referred to as rab escort protein-1 (REP-1), forms part of the rab GGTase, which attaches the geranylgeranyl isoprenoid group to certain members of the Rab family of small GTPases (Seabra *et al.*, 1992). It has been suggested that subunit A of the holoenzyme binds to the Rab protein and presents it to subunit B - which transfers the geranylgeranyl moiety - and then delivers the Rab to its target membrane (Alexandrov *et al.*, 1994). The Rab proteins are involved in intracellular trafficking and prenylation is required for correct membrane targeting. Another subunit, called REP-2, has been identified which is homologous to REP-1 and has a partially overlapping substrate specificity (Seabra *et al.*,

1995). REP-1 is ubiquitously expressed, so it is unclear why mutations do not cause systemic disease. It has been suggested that REP-2 and other homologous proteins may compensate for REP-1 in other tissues. The majority of REP-1 mutations are null (van den Hurk *et al.*, 1997), suggesting that mutations which do not completely abolish activity either fail to impair normal function or result in an extraocular phenotype.

1.6.2.iv - Prominin (mouse)-like-1 (PROML1)

The PROML1 protein is encoded by a gene which localises to 4p and is mutated in recessive retinal degeneration (Maw *et al.*, 2000). Prominin is a well conserved protein which, in the mouse, localises to plasma membrane protrusions in a range of epithelial cells. Immunocytochemistry data suggest that PROML1 accumulates in the membranous evaginations at the base of the rod outer segments (Maw *et al.*, 2000). This led to the proposal that loss of prominin impairs the formation of the evaginations and/or development of the evaginations into discs, leading to retinal degeneration.

1.6.2.v - MERTK

In the Royal College of Surgeons (RCS) rat, a model for recessive retinal degeneration, phagocytosis by the RPE of shed outer segment discs is defective, leading to photoreceptor cell death. D'Cruz *et al.* (2000) found a deletion in the gene encoding the Mertk receptor tyrosine kinase in the RCS rat. The human orthologue, *MERTK* (c-mer proto-oncogene receptor tyrosine kinase), localises to chromosomal region 2q14.1. A panel of 328 patients with a variety of retinal dystrophies were screened for mutations in this gene and three mutations were found in three RP patients (Gal *et al.*, 2000). This was the first report to show a link between defective RPE phagocytosis and human retinal disease.

1.6.3 - THE VISUAL CYCLE

1.6.3.i - RPE65

RPE65 is a microsomal protein found in large amounts in the RPE (Hamel *et al.*, 1993). The *RPE65* gene was mapped to 1p31 (Hamel *et al.*, 1994) and mutations in it were subsequently found in patients with LCA (Marlhens *et al.*, 1997), ARRP (Morimura *et al.*, 1998) and autosomal recessive childhood-onset severe retinal dystrophy (Gu, *et al.*, 1997). *RPE65*

knockout mice display an absence of rhodopsin and rod-mediated light responses, although they retain cone-mediated responses (Redmond *et al.*, 1998). In the RPE of these mice there is an accumulation of all-*trans* retinyl esters but a lack of 11-*cis* retinyl esters, indicating that the all-*trans* to 11-*cis* isomerisation step is blocked. Van Hooser *et al* (2000) orally administered 9-*cis* retinal to *RPE* (-/-) mice in an attempt to relieve the effects of this block. Commercially available 9-*cis* retinal was used because it is thermodynamically more stable than 11-*cis* retinal, for which it is a functional substitute. They observed the formation of rod photopigment and a dramatic improvement in rod physiology within 48 hours, suggesting that such a strategy may have the potential to lessen the effects of *RPE65* mutations in LCA patients.

1.6.3.ii - ABCA4 (ABCR)

Stargardt disease is the most common autosomal recessive macular dystrophy, affecting an estimated 1 in 10,000 people (Blacharski, 1988). Mutations in the *ABCA4* (*ABCR*) gene, located in 1p21-p22, are found in Stargardt disease patients (Allikmets *et al.*, 1997a). This gene encodes a member of the ATP-binding cassette (ABC) superfamily, which includes transmembrane proteins implicated in the active transport of a variety of substrates across membranes. Mutations in genes encoding members of this family result in diseases associated with defective transport (e.g. *CFTR* in cystic fibrosis). The ABCA4 protein localises to the rim of rod and cone outer segment disc membranes (Sun and Nathans, 1997; Molday, Rabin and Molday, 2000). In *ABCA4* knockout mice there is a reduction in all-*trans* retinol and an increase in all-*trans* retinal in the retina after exposure to light suggesting an accumulation of all-*trans* retinal within the disc lumen or membranes (Weng *et al.*, 1999). This has led to the suggestion that the role of ABCA4 is to transport all-*trans* retinal from the disc interior through the disc membrane (see section 1.5).

Stargardt disease patients are usually compound heterozygotes for *ABCA4* mutations, only one of which is predicted to be a complete null. Patients with two (predicted) null mutations in *ABCA4* suffer from autosomal recessive RP (Martinez-Mir *et al.*, 1998). This implies that Stargardt disease may be a partial loss-of-function phenotype while the null phenotype is RP. Since approximately 2% of the human population carry an *ABCA4* allele conferring partial or complete loss-of-function, it would be interesting to know whether such carriers develop a retinal disease, particularly age-related macular degeneration (AMD). Two groups recently addressed this question but reported conflicting findings (Allikmets *et al.*, 1997b; Stone *et al.*, 1998). Another study however, found either of two

heterozygous *ABCA4* mutations in 4% of AMD patients but only 0.9% of control individuals (Allikmets *et al.*, 1999). This would imply that carriers of these specific *ABCA4* mutations are at a higher risk for AMD, although data from ongoing studies will be important in resolving this issue.

1.6.3.iii - CRalBP

The gene encoding the cellular retinaldehyde binding protein (CRalBP) localises to 15q26 and is mutated in recessive RP (Maw *et al.*, 1997), recessive Bothnia dystrophy (Burstedt *et al.*, 1999) and recessive retinitis punctata albescens (Morimura, Berson and Dryja, 1999). The 36 kDa protein is found in Muller cells and RPE and preferentially binds 11-*cis* retinol. *In vitro*, CRalBP promotes 11-*cis* retinol oxidation to form 11-*cis* retinal (the final step in the recycling of the rhodopsin chromophore) and inhibits the esterification of 11-*cis* retinol to form the stored retinol ester (Saari, *et al.*, 1994). Recombinant CRalBP bearing a disease-causing mutation fails to bind 11-*cis* retinal. This suggests that, *in vivo*, defective CRalBP is unable to complete the regeneration of rhodopsin (Maw *et al.*, 1997). CralBP deficient mice demonstrated a more than 10-fold delay in rhodopsin regeneration, 11-*cis* retinal production and dark adaptation after illumination (Saari *et al.*, 2001). During the delay, the accumulation of all-*trans* retinyl esters was observed, indicating a block at the all-*trans* to 11-*cis* retinal isomerization step.

1.6.3.iv - 11-*cis* Retinol Dehydrogenase

The enzyme 11-*cis* retinol dehydrogenase catalyses the conversion of 11-*cis* retinol to 11-*cis* retinal (Simon *et al.*, 1995) and is found in abundance in the RPE. The *RDH5* gene encoding this enzyme maps to 12q13-q14 and is mutated in patients suffering from fundus albipunctatus. This is a rare form of autosomal recessive stationary night blindness in which rod and cone photopigment regeneration is delayed (Yamamoto *et al.*, 1999). The fact that both rods and cones are affected in this disease suggests that the same retinol dehydrogenase processes 11-*cis* retinal for both photoreceptor types.

1.6.3.v - RPE-Retinal G Protein-Coupled Receptor

Mutations in the gene for RPE-retinal G protein-coupled receptor (RGR), located at 10q23, have been found in patients with recessive RP and dominant choroidal sclerosis (Morimura

et al., 1999). This gene is abundantly expressed in the RPE and Muller cells and is related to opsin and other GPCR family members (Jiang, Pandey and Fong, 1993). The protein, which has seven transmembrane segments, preferentially binds all-*trans* retinal, in contrast to rhodopsin, which binds 11-*cis* retinal. In mammalian RGR, it is proposed that light stimulates the conversion of all-*trans* retinal to 11-*cis* retinal; this contrasts with rhodopsin, where the reverse isomerisation reaction occurs.

1.6.4 - TRANSCRIPTION FACTORS

1.6.4.i - Cone Rod Homeobox (CRX)

The *CRX* gene (for 'cone-rod homeobox-containing gene') maps to 19q13.3 and is mutated in LCA (Freund *et al.*, 1998; Sohocki *et al.*, 1998), dominant RP (Jacobson *et al.*, 1998) and dominant cone-rod dystrophy (CRD) (Freund *et al.*, 1997). CRD is characterised by an early deterioration in colour vision and visual acuity, followed by night blindness and loss of peripheral vision. The CRX protein is a transcription factor involved in the maintenance and differentiation of photoreceptor cells, stimulating the expression of retinal genes such as opsins, arrestin and interphotoreceptor retinoid-binding protein (Furukawa, Morrow and Cepko, 1997). The variety of phenotypes associated with *CRX* mutations suggests a diversity of downstream effects. These are predominantly dominant disorders (although *CRX* mutations are assocaited with recessive LCA) and it is not yet clear whether the phenotypes are due to haploinsufficiency or a dominant negative effect.

1.6.4.ii - Neural Retina Leucine Zipper

The *Neural retina leucine zipper* (*NRL*) gene maps to 14q11.1-q11.2 and encodes a retina-specific transcription factor (Yang-Feng and Swaroop, 1992). NRL promotes the transcription of rhodopsin and other retinal genes, and functions both alone and synergistically with CRX (Chen *et al.*, 1997). A T-to-A transversion at nucleotide 1942 of the *NRL* gene, which results in a ser-to-thr substitution at codon 50 of the NRL protein, was identified in all affected members of a large autosomal dominant RP family (Bessant *et al.*, 1999). In combination with CRX, the mutant NRL protein was found to promote transcription from the rhodopsin promoter more efficiently than the wild-type protein *in vitro* (Bessant *et al.*, 1999). It is possible that this effect could result in a retinopathy since

transgenic mice over-expressing rhodopsin are associated with photoreceptor degeneration (Olsson *et al.*, 1992).

1.6.5 - METABOLISM

1.6.5.i - Ornithine Aminotransferase

Gyrate atrophy (GA) is an autosomal recessive disorder resembling RP, which is characterised by night blindness, loss of peripheral vision, a decline in visual acuity and sharply demarcated circular areas of chorioretinal atrophy (Weleber and Kennaway, 1988). The disorder is caused by mutations in the gene encoding ornithine aminotransferase (OAT), which maps to 10q26 (Valle and Simell, 1995). OAT is found in mitochondria in all tissues and links glutamate, proline and ornithine metabolism. In the mouse model of gyrate atrophy, disruption of the murine OAT gene causes early morphological damage in the RPE and a slow progressive photoreceptor degeneration (Wang *et al.*, 1996). The authors subsequently showed that long-term dietary changes in the mouse model alleviate the effects of the gene disruption (Wang *et al.*, 2000). They used an arginine-restricted diet to lower ornithine levels and observed a complete absence of retinal degeneration in *oat* (-/-) mice. This demonstrates that the accumulation of ornithine is required for the retinal degeneration in GA and that effective treatment does not require the restoration of retinal OAT activity.

1.6.5.ii - PEX1, PEX2 and PAH

Mutations in the *PEX1* gene cause the infantile form of Refsum disease, a recessive peroxisome biogenesis disorder (PBD), which is characterised by RP, peripheral neuropathy, hearing defects, cerebellar ataxia, dysmorphic features and mental retardation associated with raised plasma phytanic acid (Reuber *et al.*, 1997). *PEX1* is localised to 7q21-q22 and is the human orthologue of a yeast gene that is required for peroxisomal matrix import. PBDs are caused by disrupted peroxisome function and aberrant peroxisome assembly. Shimozawa *et al.* (1999) identified *PEX2* mutations in a patient with infantile Refsum disease. Mutations in this gene were shown *in vitro* to deleteriously affect the peroxisomal protein import machinery and other peroxisome activities.

Adult Refsum disease is also associated with elevated plasma phytanic acid levels and results from mutations in the gene encoding phytanoyl-CoA hydoxylase (PAH), which maps to 10p15.3-p12.2 (Jansen *et al.*, 1997; Mihalik *et al.*, 1997). Other charactersitics of

the disease are RP, peripheral polyneuropathy, cerebellar ataxia, high CSF protein (but no dysmorphic features or mental retardation) with onset varying between early childhood and the fifth decade. Phytanoyl-CoA hydroxylase is also a peroxisomal enzyme and catalyses the initial step in phytanic acid α-oxidation. Elimination of phytanic acid from the diet results in neurological improvement and prevents progression of the retinal dystrophy.

1.6.6 - MISCELLANEOUS

There is a growing group of retinal degenerative disorders associated with mutations in genes with miscellaneous functions; these are described below. In others, there is little functional information available regarding the gene product; these have been entered into Table 1.1 at the end of this section.

1.6.6.i - Usherin (USH2A)

Usher syndrome type IIa is caused by mutations in the *usherin* (*USH2A*) gene located in 1q41 (Eudy *et al.*, 1998). Mutations in this gene account for 95% of type 2 Usher syndrome patients. The gene encodes a predicted 171.5 kDa product containing laminin epidermal growth factor and fibronectin type III motifs. These motifs tend to be found in cell adhesion molecules and in components of the extracellular matrix and basal lamina. Bruch's membrane, the interphotoreceptor cell matrix and the cochlea are all rich in extracellular matrix proteins.

1.6.6.ii - Epidermal Growth Factor (EGF)-Containing Fibrillin-Like Extracellular Matrix Protein-1 (EFEMP1)

Dominant Doyne honeycomb retinal degeneration (DHRD, clinically similar to Malattia Leventinese) is an autosomal dominant disorder that closely resembles AMD. Approximately 50% of registered blindness in the Western world is due to AMD (Stone *et al.*, 1999). DHRD is characterised by the presence of yellow/white deposits (drusen) beneath the RPE and a progressive central vision loss. An R345W mutation was found in the *EFEMP1* gene in 5 DHRD families (Stone *et al.*, 1999). This widely expressed gene maps to 2p16-p21 and is homologous to members of the fibulin extracellular matrix glycoprotein family.

1.6.6.iii - TIMP3

Sorsby's fundus dystrophy (SFD) is a rare autosomal dominant retinal degeneration. The characteristics of this disorder include a slow recovery following exposure to bright light and a loss of central vision. The gene encoding tissue inhibitor of metalloproteinase-3 (TIMP-3), which maps to 22q12.1-q13.2, is mutated in this disease (Weber *et al.*, 1994). TIMPs act as inhibitors of the matrix metalloproteinases that function in the degradation of the extracellular matrix (Apte, Mattei and Olsen, 1994). In SFD there is an abnormal accumulation of a lipid-rich material between the choroid and the photoreceptor cell layers. It was initially suggested that this arises because of an imbalance in the breakdown and synthesis of the extracellular matrix (caused by the *TIMP-3* mutation) and prevents the normal supply of nutrients (including Vitamin A) from the choroidal blood supply to the photoreceptors. In testing this hypothesis Jacobson *et al.* (1995) found that high doses of oral vitamin A (reversibly) corrected the reduced rod sensitivity in patients at early stages of the disease. A recent report however, showed that mutant *TIMP-3* genes are expressed and that the proteins retain the ability to inhibit metalloproteinases (Langton *et al.*, 2000). Unlike wild-type TIMP-3 however, all of the mutant proteins studied formed dimers. This, together with the recent observation that *TIMP-3* expression is increased in SFD (Chong *et al.*, 2000) suggests that accumulated, dimerised mutant TIMP-3 plays an active role in the disease process.

1.6.6.iv - TULP1

In *tubby* mice, a recessive mutation in the *tub* gene leads to retinal degeneration, cochlear deafness, obesity and insulin resistance (Kleyn *et al.*, 1996). In humans, *TUB* and *Tubby-like proteins 1* (*TULP1*) and *2* (*TULP2*) are homologous genes with a highly conserved C-terminal domain (North *et al.*, 1997). *TULP1* is expressed solely in the retina and maps to 6p21.3 so was considered a candidate gene for a recessive RP locus (*RP14*) which had been mapped to this chromosomal location by linkage analysis. *TULP1* mutations were subsequently found in recessive RP patients (Banerjee *et al.*, 1998; Hagstrom *et al.*, 1998; Gu *et al.*, 1998). *Tulp1(-/-)* mice display a retinal degeneration of early-onset, with a progressive, rapid photoreceptor loss (Ikeda *et al.*, 2000) and substantial extracellular vesicle accumulation around the distal inner segments (Hagstrom *et al.*, 1999). The results of experiments carried out to characterise the mouse tubby protein indicate that the tubby-like proteins belong to a unique category of bipartite transcription factors (Boggon *et al.*, 1999).

1.6.6.v - Tocopherol Transfer Protein Alpha (TTPA)

TTPA is an intracellular vitamin E (α-tocopherol) carrier protein that is structurally similar to CRalBP. The primary antioxidant in photoreceptor outer segments is vitamin E, which is presumed to prevent photo-oxidation damage. When experimental animals are deprived of vitamin E they develop RP (Robinson, Kuwabara and Bieri, 1982). Two patients with autosomal recessive RP and low serum vitamin E were studied and found to carry His101Gln mutations in the *TTPA* gene (Yokota *et al.*, 1996). Mutations in this gene, which maps to 8q13.1-q13.3, are also associated with ataxia with isolated vitamin E deficiency (Ouahchi *et al.*, 1995).

1.6.6.vi - Crumbs Homologue 1 (Crb1)

The human homologue of the *Drosophila melanogaster* crumbs (CRB) protein is encoded by a gene, *crumbs homologue 1* (*CRB1*), which is mutated in a severe form of autosomal recessive RP called RP with para-arteriolar preservation of the RPE (see section 1.7.2) (Den Hollander *et al.*, 1999). The *CRB1* cDNA was cloned using a subtractive hybridisation technique in an attempt to identify genes expressed in the retina and RPE. The *CRB1* gene mapped to chromosomal region 1q31-q32.1, a region implicated in RP. On the basis of the homology to the *Drosophila* protein a role has been proposed for CRB1 in the maintenance of cell polarity in the retina and in cell-cell interaction.

CRB1 mutations have also been found to result in Leber congenital amaurosis (Lotery *et al.*, 2001). Mutations were detected in 21 patients out of a cohort of 233, amounting to 9% of the total.

Table 1.1: Miscellaneous genes mutated in retinal degenerative diseases.

Disorder	Gene/Protein	Location	References
Abetalipoproteinemia	Microsomal triglyceride transfer protein	4q24	Sharp *et al.*, 1993; Shoulders *et al.*, 1993.
Bardet-Biedl syndrome	*MKKS* (McKusick-Kaufman syndrome), protein is putative chaperonin molecule	20p12	Slavotinek *et al.*, 2000.
Best disease (vitelline macular dystrophy)	Bestrophin	11q13	Marquardt *et al.*, 1998.
Dominant RP (RP1)	Oxygen-regulated photoreceptor-specific protein/ RP1 protein	8q11-q13	Pierce *et al.*, 1999; Sullivan *et al.*, 1999.
Enhanced S-cone syndrome (ESCS)	*Nuclear receptor subfamily 2, group E, member 3 (NR2E3) gene*/ photoreceptor-specific nuclear receptor (pnr) protein	15q23	Haider *et al.*, 2000.
LCA type 4	Arylhydrocarbon-Interacting Receptor Protein-Like 1 (AIPL1)	17p13.1	Sohocki *et al.*, 2000.
Recessive Batten disease	*CLN3* gene/ Batten disease protein	16p12.1	International Batten Disease Consortium, 1995.
RP with myopathy	PGK1, phosphoglycerate kinase	Xq13.3	Tonin *et al.*, 1993.
Spinocerebellar ataxia 7 with pigmentary macular degeneration	SCA7	3p21.1-p12	David *et al.*, 1997.
Usher syndrome type 1C	*USH1C*/Harmonin, a PDZ-domain containing protein	11p14.3	Verpy *et al.*, 2000; Bitner-Glindzicz *et al.*, 2000.
Usher syndrome type 1D	Novel cadherin gene *CDH23*	10q21-q22	Bork *et al.*, 2001; Bolz *et al.*, 2001.
X-linked juvenile retinoschisis	X-linked retinoschisis-1 protein (RS1)	Xp22.2	Sauer *et al.*, 1997.
X-linked RP (RP2)	*RP2* gene, RP2 protein	Xp11.3	Schwahn *et al.*, 1998.

1.7 - RETINITIS PIGMENTOSA

1.7.1 - Symptoms of Retinitis Pigmentosa

There is wide clinical variation among RP patients but common to all sufferers are night blindness (nyctalopia) and a progressive visual field loss. This leads to a severe visual disability and can include loss of central vision. The combination of peripheral and central visual loss is particularly debilitating. Many patients aged between 60 and 70 have little or no functional vision and a little under half of all RP patients develop cataracts, usually in the later stages of the disease. RP is generally caused by a primary or secondary degeneration of rod photoreceptor cells, but cone photoreceptors also degenerate as the disease progresses. Rod cell death starts at the mid-periphery of the retina, the most rod-dense region, but then spreads to involve the far periphery and macula. Rod cells respond to dim light and are found in all areas of the retina except the fovea centralis, and the nyctalopia and visual field contraction characteristic of RP reflect this. Towards the end of the disease process, photoreceptor cell degeneration is accompanied by reactive changes in the RPE and Muller cells, atrophy of the retinal vasculature and death of inner retinal neurons (Milam and Li, 1995).

Ophthalmic examination of RP patients reveals varying degrees of depigmentation or atrophy of the RPE, pigment deposition in the retina and narrowing of the retinal arterioles (secondary to retinal atrophy) (Deutmann, 1977). Pigment-laden cells from the RPE migrate to the mid-periphery of the retina. The pigment deposits appear as black clumps and strands. They are found most prominently in the periphery, and frequently, pigment deposition within blood vessel walls produces a perivascular pattern, the so-called 'bone spicule pigment'. Small, irregular clumps and spots of pigment may be seen almost as often and, occasionally, patients have a mixture of clumped pigment and bone spicule pigment. In many cases, regions of the retinal fundus (the posterior aspect of the eyeball, visible through the pupil with an ophthalmoscope) not involved in pigment deposition are described as appearing "moth eaten", or as having a "salt and pepper" pigment distribution. Figures 1.8 and 1.9 show fundus photographs from unaffected and RP-affected individuals respectively. Nearly all RP patients go through a stage in which there is little or no visible pigment deposition. This stage persists in some patients and their condition was formerly described as retinitis pigmentosa sine pigmento, to demarcate an atypical type of the RP. However, the majority of these patients have since been found to produce an electroretinogram (ERG) pattern indicative of the cone-rod degeneration form of RP (see section 1.7.2).

A scotopic ERG measures the electrical activity in a dark-adapted retina in response to flashes of light and can be used to assess rod photoreceptor function (the photopic ERG is performed in the light-adapted state and measures the cone-mediated response). Scotopic ERGs typically show reduced or absent amplitudes in RP patients and may also be delayed temporally (Heckenlively, 1988a). Occasionally, it is possible to identify ERG changes in presymptomatic cases of RP (Humphries, Kenna and Farrar, 1992).

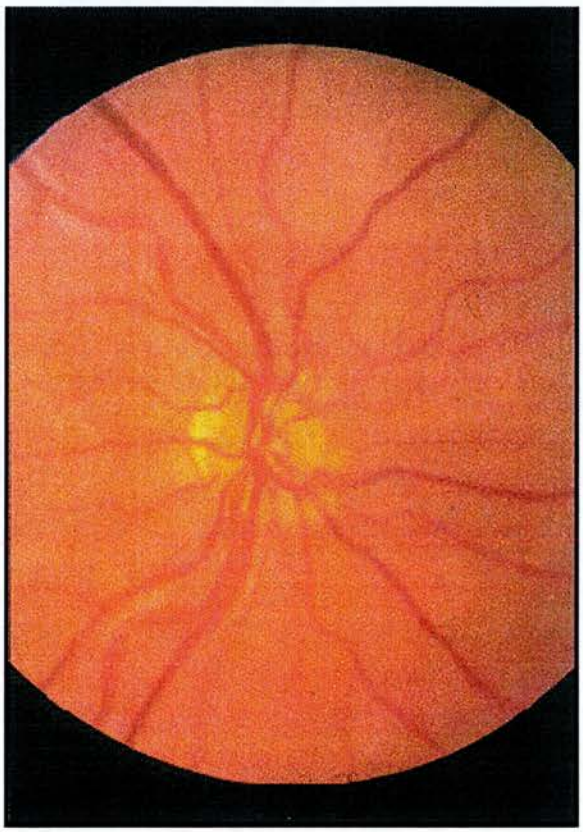


Figure 1.8: Photograph of a fundus from an unaffected individual.

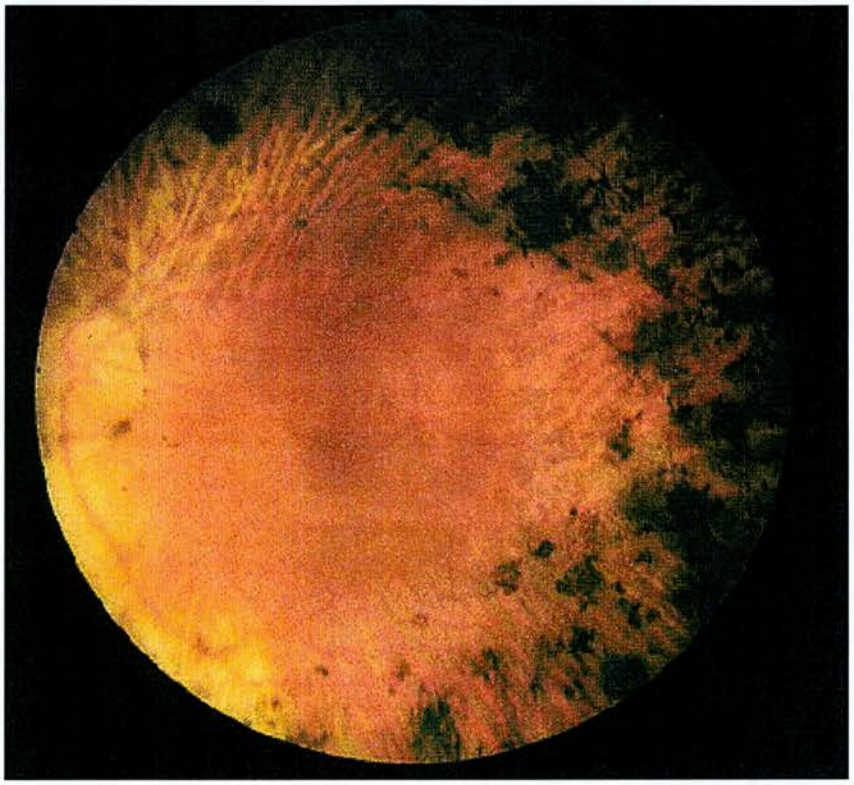


Figure 1.9: Photograph of a fundus from an RP patient.

As RP progresses and rod photoreceptor cells continue to degenerate, the optic nerve head develops a waxy, pale appearance, the retinal blood supply declines following blood vessel attenuation and the retina becomes thinner in appearance. Staining retinal sections from RP patients with anti-rhodopsin antibodies revealed that rod and cone photoreceptor cell death is preceded by a shortening of the outer segments (Milam and Li, 1995). In most patients, the macula, and central vision, is spared until relatively late in the course of the disease. When central vision does eventually deteriorate it is frequently due to macular involvement, macular cysts or posterior subcapsular cataracts, which are characterised by yellow crystalline changes in the lens (Heckenlively, 1988b). In addition, it is thought that dying or defective rod photoreceptors induces cell death in adjacent cones (Remé *et al.*, 1998).

1.7.2 - Subtypes of Retinitis Pigmentosa

RP can be divided into two large subgroups: primary RP, where the disease process affects only the eyes, and secondary RP, where the retinal degeneration accompanies a single or multiple organ system disease (Heckenlively, 1988c). Usher syndrome and Bardet-Biedl syndrome are the most common causes of secondary RP (Wright *et al.*, 1995). Primary RP is, in turn, subdivided into rod-cone degenerations, cone-rod degenerations and those with congenital onset (Heckenlively 1988c). Rod-cone and cone-rod degenerations are discriminated on the basis of the ERG. In typical RP, rods are involved to a greater degree than cones and the rod-mediated (scotopic) ERG is more affected than the cone-mediated (photopic) ERG: this is rod-cone degeneration. Some patients have a progressive visual field loss with absent or late-onset night blindness and a more severely affected cone-mediated ERG: this is cone-rod degeneration. The cone-rod and rod-cone degenerations can be inherited in autosomal dominant, autosomal recessive or X-linked recessive fashion. It is generally accepted that X-linked and autosomal recessive RP are usually more severe than autosomal dominant RP (Bird and Jay, 1994). Cone-rod and rod-cone degenerations can present as simplex or multiplex cases. If there is no family history of RP and the patient is the only affected member of a pedigree the term *simplex* is used. The term *sporadic* is appropriate only when an environmental or new mutational event is suspected to be the cause of the disease. This may be the case following exposure to chloroquine (used in the treatment of arthritis or malaria), or thioridazine (used to treat psychoses), or as a result of rubella or cytomegalovirus infection (Heckenlively, 1988a). When two or more siblings but no other members of the family are affected, the term *multiplex* is used. Although multiplex sibships do not all follow the same mode of inheritance, most are considered likely to be

autosomal recessive (Heckenlively, 1988b).

Two broad clinical subdivisions of autosomal dominant RP exist: type I, a so-called "diffuse" form, and type II, a "regional" form (Massof and Finkelstein, 1981; Lyness *et al*, 1985). Type I is characterised by a childhood-onset night-blindness, widespread degeneration of rods and delayed cone involvement. In type II there is an adult-onset night-blindness and regionalised loss of rod and cone function. An additional variant of autosomal dominant RP is sector RP, where the retinopathy is restricted to the inferior nasal quadrant (and less commonly the inferior temporal quadrant). Sector RP can be inherited by all modes but the autosomal dominant condition is the most common (Heckenlively, 1988d).

ARRP has been subdivided into two broad groups by Kaplan *et al.* (1990) on the basis of age of onset: precocious onset with severe progression (mean age of onset of 7.5 years), and later onset with mild progression (mean age of onset 17 years). Grondahl (1987a) and Bonneau *et al.* (1992) also described a "senile" form of ARRP, where vision is unimpaired until the fifth or sixth decade. Subtypes of ARRP include retinitis punctata albescens (in some cases due to mutations affecting rhodopsin (Souied *et al.*, 1996) or CRalBP (Morimura, Berson and Dryja, 1999)) and preserved para-arteriolar retinal pigment epithelium (PPRPE; due to *CRB1* mutations, see section 1.6.6.vi), both of which are rod-cone degenerations. Retinitis punctata albescens is distinct from other forms of RP and is characterised by a fundus marked with numerous punctate spots of a whitish-yellow colour (Ellis and Heckenlively, 1988). PPRPE is a severe, childhood onset RP variant in which there is some preservation of the RPE around blood vessels (Ellis and Heckenlively, 1988).

Choroideremia is another example of a rod-cone dystrophy. This is an X-linked, progressive degeneration, in which the choroid and retina ultimately become atrophic (Forrester *et al.*, 1995). Among the different types of congenital RP are Leber congenital amaurosis (typical form) and congenital RP with macular coloboma.

1.7.3 - Prevalence of Retinitis Pigmentosa

Retinitis Pigmentosa is a major cause of blindness and a number of studies have attempted to determine its frequency in the general population. These estimates range from 1 in 1333 in Birmingham, UK (Bundey and Crews, 1984a) to 1 in 7000 in Switzerland (Ammann *et al.*, 1965) but a figure close to 1 in 4000 is generally accepted. Table 1.2 presents the results of several of these studies.

Table 1.2: Prevalence of Retinitis Pigmentosa

COUNTRY	PREVALENCE	REFERENCE
China	1 in 4016	Hu (1987).
Israel	1 in 4500	Merin and Auerbach (1976).
North America	1 in 3700	Boughman *et al.* (1980).
Norway	1 in 4440	Grondahl (1987b).
Sardinia	1 in 3355	Fossarello *et al.* (1993).
Switzerland	1 in 7000	Ammann *et al.* (1965).
UK	1 in 1333-1666	Bundey and Crews (1984a).
USA	1 in 4756	Bunker *et al.* (1984).

RP is present in all ethnic groups, and tends to affect similar percentages of the population. Efforts have been made to determine the percentage of RP patients in each genetic subgroup and the results of several studies are summarised in Table 1.3. The relative proportions vary from study to study. This may be attributable to differences in the methods used by the researchers, different sample sizes or differences in the populations being studied, such as levels of consanguinity or sex and age distribution. For example, consanguineous marriages are more common in the Sardinian population than in others and this may lead to greater numbers of patients with autosomal recessive RP.

Some of these studies suffer from the authors' reliance on questionnaires to collect information about diagnoses and family histories. Boughman *et al.* (1980) used this method exclusively to determine the frequency of each inheritance pattern for RP in America. They counted the number of autosomal dominant and X-linked RP patients and then assigned those remaining to the autosomal recessive and sporadic categories. They did not attempt to actively identify and count ARRP cases, to confirm the diagnoses clinically or to separate these as a group from isolated cases. Several of the ARRP cases may have been ADRP: if penetrance was lower than the authors presume, as suggested by Bundey and Crews (1984b), then it is possible that some apparently unaffected relatives of isolated RP patients carried RP alleles. These cases may have been detected had family members been examined so that asymptomatic or presymptomatic signs of the disease could have been observed. Similarly, some of the sporadic cases may have been XLRP. By examining female relatives of isolated

male patients, carrier females could have been detected and these families correctly classified. Macrae (1982) studied 698 Canadian RP patients (in 274 families), 130 (in 110 families) by examination and the rest by questionnaire. When the mode of inheritance could not be established relatives were not examined. A high proportion of simplex RP was reported, some of which could possibly have been assigned to other classifications had relatives been examined. Bunker *et al.* (1984) made great efforts to identify all the RP patients within the state of Maine, USA. Information was collected through patient interviews and, when the mode of inheritance was unclear, examinations were made. In the case of isolated males, the mother or daughter was examined for evidence of XLRP carrier status.

Table 1.3: Proportions of RP according to mode of inheritance. Numbers indicate percentages of patients in each category, percentages of families in parentheses (where information is available).

Country	ARRP (%)	ADRP (%)	XLRP (%)	Simplex (%)	Multiplex (%)	Other (%)	No. cases (families)	Reference
Canada	(28)	(16)	(8)	(48)	-	-	698 (274)	Macrae 1982.
Denmark	19.0 (17.9)	8.4 (4.4)	14.3 (6.2)	42.9 (56.7)	11.2 (9.3)	4.2 (5.5) **	837 (633)	Haim 1992.
England	(40.8)	(24.4)	(22.3)	(11.6)		(0.9) †	1154 (426)	Jay 1982 ('GC' series)
England	11.7	25.0 (22.5)	15.8	42.5	-	5.0††	120 (113)	Bundey and Crews 1984b.
USA	19 (19)	19 (26)	8 (16)	52 (38)		2 (1)**	173 (124)	Fishman 1978
USA	84*	10	6	-	-	-	670 (648)	Boughman *et al.*, 1980.
USA	20 (19)	43 (19)	8 (8)	23 (46)	6 (8)	-	168 (85)	Bunker *et al.*, 1984.
USA	16 (16)	28 (22)	12 (9)	41 (50)	-	3 (3)**	370 (300)	Boughman and Fishman 1983.

*This category described as "autosomal recessive and sporadic" by the authors; ** uncertain classification; † adopted; †† includes blind-registered XLRP carrier female (0.7%) and unclassifiable (3.6%), 18 patients (13% of families) with autosomal recessive syndromes have been removed from this category.

One of the most thorough studies was reported by Bundey and Crews (1984b). They investigated 138 cases of RP within the city of Birmingham, UK, examining a large number of patients and their relatives (18 autosomal recessive RP syndrome cases were removed from this set for comparison with the other data in table 1.3). Of 30 individuals with ADRP, eight had asymptomatic parents and the correct mode of inheritance was only determined

after examining these family members. Without this practice these eight (21% of the ADRP cases) would have been added to the simplex group. Seven out of 19 (37%) XLRP patients were classified in the same way. As a result of examining family members, 25 cases were removed from the initial group of 76 isolated patients and reassigned to the ADRP, ARRP and XLRP categories. The remaining 51 cases comprised 26 females and 25 males, so an excess of unrecognised XLRP is unlikely to have been concealed within this figure. Bundey and Crews also reported one of the lowest incidences of ARRP (11.7% of the total) of the studies discussed here. Moreover, half of these ARRP patients had consanguineous parents, indicating that, in Birmingham (UK) at least, this is a rare disease and individual disease gene frequencies are low. The results of this study are presented separately in Figure 1.10. Unlike Bundey and Crews (1984b), Jay (1982) estimated the proportion of simplex cases that will have been inherited in an autosomal recessive manner and added these to the ARRP category (Table 1.3). Jay consequently reported a much larger proportion of ARRP than Bundey and Crews (1984b) despite both studying RP in UK cities.

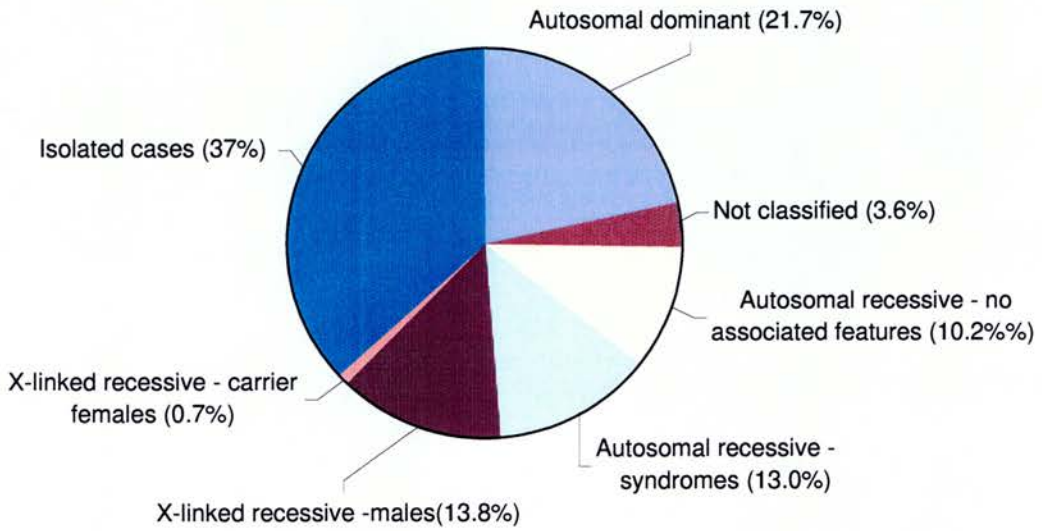


Figure 1.10: Genetic categories of RP patients resident in Birmingham in 1978 (Bundey and Crews 1984b).

Ascertainment bias may also alter the proportion of each genetic subtype. Bundey and Crews (1984b) showed that the relative proportions of each of the genetic subtypes of RP vary according to the mode of ascertainment. This implies that the results of some studies will suffer if they rely on a single mode of ascertainment, whereas the authors of other studies, such as Bundey and Crews (1984), Haim (1992) and Bunker *et al.* (1984) have attempted to minimise such errors.

Some of the differences between the various studies will be real and not due to sampling error, ascertainment bias or incorrect assignment. Genuine differences between studies reflect population differences (see the studies of Swiss RP by Ammann *et al.* (1965) and Sardinian RP by Fossarello *et al.* (1993)). What none of these studies can accurately measure is the proportion of isolated patients with non-genetic RP. Jay (1982) attempted to estimate this by segregation analysis and concluded that as many as 19.4% of simplex cases may be non-genetic. Between different population groups, environmental differences may lead to a greater or lesser contribution of phenocopies and novel mutations to the isolated classification.

1.7.4 - X-linked RP-affected males (hemizygotes)

RP has been found to affect more men than women, first by Nettleship in 1908 when he established that 61.21% of RP patients were male (Nettleship, 1908), and again, in 1966, when Sorsby concluded that 59.5% of people who were blind due to RP or allied disorders were male (Sorsby, 1966). This sex-bias may arise either because of ascertainment bias (at the time of early studies, men who experienced visual difficulties in the workplace may have been more likely to have been ascertained than women who worked in the home (Bunker *et al.*, 1984)) or due to the presence of X-linked inheritance: males with one defective X-linked recessive gene (hemizygotes) develop the disease, while females with one defective gene (carriers) generally show few or no symptoms. The average age of onset in XLRP-affected males has been estimated to be 7.2 (±1.7) years (Hussels-Maumenee *et al.*, 1975), the first common symptom being night blindness. Diagnosis may be made as early as age 4 in families associated with the disease, or as late as the middle teenage years if there are no currently affected family members. In excess of half of all XLRP males exhibit symptoms by ten years of age (Bundey and Crews 1984b). Visual loss tends to be severe in XLRP patients. During the second decade, the majority of patients develop increased visual difficulties, and by the mid-teens there is a noticeable visual field loss and a reduction in visual acuity (Bird and Heckenlively, 1998). Visual acuity continues to deteriorate in the third decade, something that usually happens later in other forms of RP, and subcapsular cataracts may develop. By the end of the fourth decade, most males are in the advanced stages of the disease and only 16% retain any useful vision (Bundey and Crews 1984b). Some patients develop nystagmus (involuntary, rapid movements of the eyeball) and, as a consequence of this, a phenomenon called oscillopsia, which is a disturbing visual sensation that stationary objects are swaying back and forth.

1.7.5 - Carrier females (heterozygotes)

The term *obligate carrier* is used to describe a woman who unambiguously carries a mutation on one X chromosome because she is the daughter or mother of an affected male. Female carriers may show no evidence of the disease, or conversely, they may appear to be as severely affected as male XLRP patients (Fishman, Weinberg, and McMahon, 1986). This variation may be explained by Lyonisation, the random inactivation of one of the two X-chromosomes early in embryogenesis, which is a feature of X-linked inheritance (Lyon, 1972). Any visual loss in carrier females tends to commence much later in life than in hemizygous patients, but once it has started, progression may be as rapid as in affected males (Bird and Heckenlively, 1988). Carrier females may display abnormal dark adaptation, reduced ERG amplitudes, visual field defects and abnormal cone flicker tests. However, most tend to suffer only a mild, late-onset retinal degeneration with few clinical symptoms.

1.8 - GENETIC MAPPING OF XLRP LOCI

Six loci on the X chromosome have been associated with XLRP: RP2 (Schwahn *et al.*, 1998), RP3 (Meindl *et al.*, 1996), RP6 (Breuer, Musarella and Swaroop, (2000), RP15 (McGuire *et al.*, 1995), RP23 (Hardcastle et al., 1999) and RP24 (Gieser *et al.*, 1998) (Figure 1.11). The vast majority of kindreds are of the RP2 or RP3 types (as indicated by heterogeneity and homogeneity analyses of large numbers of XLRP families from around the world (Ott *et al.*, 1990; Teague *et al.*, 1994; Musarella *et al.*, 1990)) and the remaining loci have each been implicated following the examination of single XLRP families (with RP15 eventually being withdrawn (Mears *et al.*, 2000), see below). RP2 accounts for an estimated 20-25% of XLRP cases, and RP3 for an estimated 70-75%. (Ott *et al.*, 1990; Teague *et al.*, 1994).

RP2 was the first XLRP locus to be genetically mapped (Bhattacharya *et al.*, 1984) and subsequent haplotype analysis placed the RP2 gene in a 5 cM interval between the markers DXS8083 (at Xp11.3) and DXS6616 (at Xp11.23) (Thiselton *et al.*, 1996). By targeting this interval using the YAC representation hybridisation technique, the gene was positionally cloned by Schwahn *et al.* (1998). The *RP2* gene is ubiquitously expressed and consists of 5 exons encoding a 350 amino acid polypeptide. The role of the protein is yet to be determined, but it shows homology to cofactor C, which plays a role in the β-tubulin folding pathway. In a subsequent study, Hardcastle *et al.* (1999) found that 6 out of 33

patients (18.2%) carried *RP2* mutations (one missense and three nonsense mutations, one 5 base pair (bp) deletion and one 1 bp insertion).

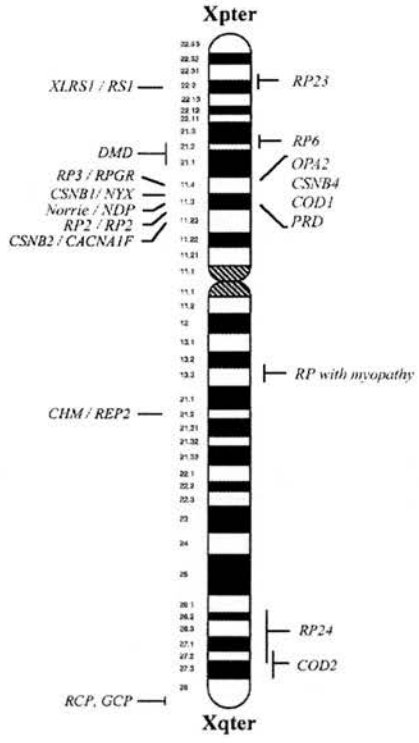


Figure 1.11: Locations of retinal degenerative disease genes on the X-chromosome (Breuer et al., 2001). XLRS1/RS1: X-linked retinoschisis/retinoschisin protein; OPA2: X-linked optic atrophy; DMD: Duchenne muscular dystrophy; CSNB1: X-linked CSNB/nyctalopin protein; Norrie/NDP: Norrie disease/Norrie disease protein; COD1: X-linked cone dystrophy; PRD: primary retinal dysplasia; CSNB2/CACNAIF: 1X-linked CSNB/L-type voltage-gated calcium channel alpha-1 subunit; CHM/REP2: choroideremia/geranylgeranyl transferase Rab escort protein 1; COD2: X-linked progressive cone dystrophy 2; GCP: protanopia and rare macular dystrophy in blue cone monochromacy/red cone opsin; RCP: deuteranopia and rare macular dystrophy in blue cone monochromacy/green cone opsin.

Statistical evidence for the RP6 locus was initially presented by Ott *et al.* in 1990 (Ott *et al.*, 1990). Later, Breuer, Musarella and Swaroop (2000) carried out genetic analysis in a single XLRP family by analysing microsatellite markers spanning the X-chromosome. Markers from RP2, RP3, RP15, RP23 and RP24 excluded these loci as responsible for the disease in this family. The workers concluded that the RP6 gene lies in a 2 cM interval between markers DXS1036 and DXS1017. It was subsequently reported, however, that the mother in this family displays an RP phenotype (Breuer *et al.*, 2001). This raises the possibility that the RP in this family is inherited in an autosomal dominant fashion rather than X-linked.

McGuire *et al.* (1995) described a new RP locus, RP15, in a family with dominant cone-rod degeneration. Microsatellite markers were used to test for linkage to the disease

locus and in this way they excluded other X-linked disease loci causing retinal degeneration. Linkage was detected for one marker however, DXS989 (LOD score 3.3), and subsequent haplotype analysis, carried out using 9 additional microsatellite markers, indicated that the disease locus was in chromosomal region Xp22.13-p22.11. However, when a female in this pedigree was identified as affected following a clinical re-evaluation, the disease was remapped to a 19.5-cM interval in Xp11.4-p21.1 that overlapped with RP3. A mutation was found in this family in RPGR exon ORF15 (see section 1.10 below) (Mears *et al.*, 2000).

Hardcastle *et al.* (2000) performed haplotype analysis (in a single XLRP family) with 34 polymorphic markers spanning the entire X chromosome. They presented linkage data describing a new locus for RP distinct from RP2 and RP3. RP23 (an atypical form of XLRP, characterised by very early onset loss of vision (at approximately 2 years of age)) was localised to the distal short arm of the X chromosome (Xp22) between markers DXS1223 and DXS7161. The authors determined that the *RS1* gene (Sauer *et al.*, 1997), which is found in this large interval, is not mutated in this family.

Gieser *et al.* (1998) mapped the RP24 locus by haplotype and linkage analysis in a single large XLRP pedigree with seven affected members. RP24 patients suffer early-onset rod dysfunction, then at later stages they have little or no rod or cone function and display the clinical hallmarks of RP. Mapping was achieved by genotyping 52 microsatellite markers spanning the entire X chromosome. The RP24 locus was narrowed down to a 23 cM interval between markers DXS8094 and DXS8043. A LOD score of 4.21 was obtained with the marker DXS8106 at Xq26-27.

1.9 - Positional cloning of the RP3 gene

Positional cloning is a strategy used to identify genes responsible for inherited disease in the absence of prior information about the location of the gene or function of the gene product. Genetic analysis of affected families is first carried out using polymorphic markers to establish the chromosomal location of the disease locus. Gross chromosomal rearrangements can aid this process by identifying a genomic region in which efforts can be concentrated. The RP2 locus was the second disease gene to be assigned an accurate chromosomal localisation using recombinant DNA methods (after Huntington disease in 1983; Gusella *et al.*, 1983) when it was linked to a polymorphic marker at Xp11.3 in 1984 (Bhattacharya *et al.*, 1984). A second XLRP locus, at Xp21, was implicated following the study of a single large XLRP kindred and the analysis of a deletion in an individual (BB) suffering from Duchenne muscular dystrophy, chronic granulomatous disease, McLeod

syndrome and RP (Nussbaum *et al.*, 1985; Francke, *et al.*,1985). RP3 was mapped to a narrow region (approximately 520 kilobases (kb) in size) flanked by the markers OTC and DXS1110 following the scrutiny of deletions in XLRP patients and recombinations identified by haplotype analysis (Musarella *et al.*, 1991; Roux *et al.*, 1993). The proximal breakpoint of the BB deletion is 40 kb centromeric to DXS1110, therefore initial attempts to find the RP3 gene focused on identifying transcribed sequences within this region. It transpired that the RP3 gene was not within the BB deletion. A patient was identified, NF, suffering from Duchenne muscular dystrophy and chronic granulomatous disease but not RP, with a deletion extending through the proximal portion of the BB deletion for several kilobases in a centromeric direction (Brown *et al.*, 1996). A deletion in another patient, SB (De Saint-Basile *et al.*, 1988), who had chronic granulomatous disease, McLeod syndrome and RP, extended 400 kb centromeric to the BB proximal deletion breakpoint (Brown *et al.*, 1996). A 75 kb deletion in another patient, MO, who suffered from classic RP, overlapped with the proximal part of the SB deletion and led to the identification of a novel retina-expressed gene (called *ETX1/SRPX*), although no causative mutations could be found (Dry *et al.*, 1995; Meindl *et al.*, 1995). *RPGR* was identified as the RP3 gene after the sequencing of 85 kb of genomic DNA from two cosmid clones spanning the MO deletion. Sequence analysis using gene recognition programs identified two expressed sequence tags (EST) and seven putative exons. Polymerase chain reaction (PCR)-amplification of the region between one of the ESTs (Accession number D20187) and the most 5' predicted exon, and exhaustive 5' and 3' primer extension and rapid amplification of cDNA ends (RACE) revealed a transcript of 2784 bp with an open reading frame of 2445 bp, encoding an 815 amino acid translation product (Meindl *et al*, 1996). Loss-of-function mutations (two small intragenic deletions, two nonsense and three missense mutations in highly conserved residues) were found in the *RPGR* gene in seven families with XLRP but not in control chromosomes.

1.10 - RPGR

RPGR appears to be ubiquitously expressed: hybridising an *RPGR* probe to a northern blot revealed the presence of a 2.9 kb band in all tissues examined except, paradoxically, retina. However, *RPGR* was detected in retina and RPE cDNA libraries, and foetal RPE cells by RT-PCR. The *RPGR* gene was reported to consist of 19 exons, covering 60 kb of genomic sequence (Meindl *et al.*, 1996; Roepman *et al.*, 1996).

The outstanding feature of the predicted RPGR protein is the homology of its N-terminal half to the Regulator of Chromosome Condensation 1 (RCC1) protein. The RCC1-

like domain of RPGR extends from amino acid number 39 (corresponding to the middle of exon 2) to amino acid 365 (corresponding to the middle of exon 10). RCC1 consists of seven tandem repeat units, 51-68 residues long, and is a guanine nucleotide exchange factor (GEF) for the small nuclear GTPase Ran, which is concerned with nucleocytoplasmic transport and cell cycle control (Ohtsubo *et al.*, 1987; Sazer and Dasso, 2000). An RCC1-like domain occurs in several other proteins, including p532 (Rosa *et al.*, 1996), which appears to act as a GEF for Rab and Arf GTPases, suggesting that its presence indicates GEF activity.

The crystal structure of RCC1 was solved and found to have a pseudo-sevenfold symmetry and the appearance of a seven-bladed propeller (Renault *et al.*, 1998; Figure 1.12).

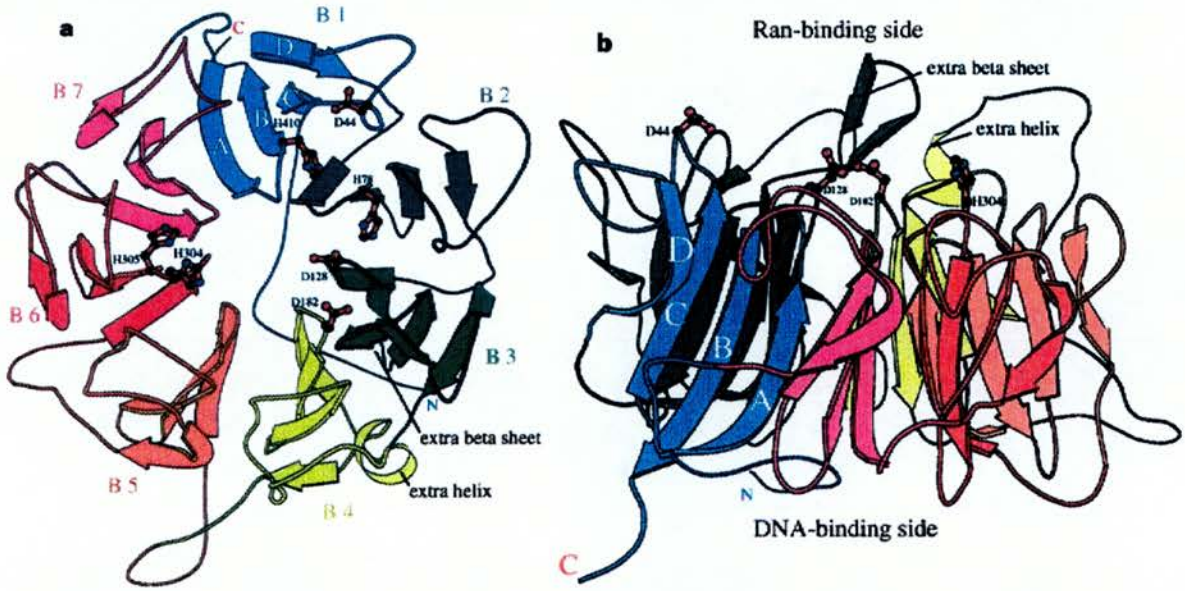


Figure 1.12: Ribbon diagram of the seven-bladed propeller structure of RCC1 viewed along (a) or perpendicular (b) to the shaft through the centre (from Renault *et al.*, 1998).

The propeller blades are each formed by a four-stranded antiparallel β-sheet, the four strands being referred to as A, B, C and D. This propeller structure is similar to that of the β-subunit of the heterotrimeric G proteins and some enzymes (Lambright, *et al.*, 1996). G-protein β-subunits contain WD40 repeats; these ~40-residue configurations terminate in conserved Trp-Asp (WD) motifs and are distinct from RCC1 repeats (Neer, *et al.*, 1994). The seven RCC1 sequence repeats do not coincide directly with the seven structural repeats, but are one-half of a repeat out of step with each other. The first half of each sequence repeat encodes strands C and D of the corresponding structural repeat (propeller blade), and the second half of the sequence repeat encodes the A and B strands of the next structural repeat. For example, the first half of the third sequence repeat encodes strands C and D of blade 3, and the second half of the third sequence repeat encodes strands A and B of blade 4, and so

on. Strands C and D of the first repeat are encoded by the first half of the first sequence repeat, and strands A and B are encoded by the second half of the seventh sequence repeat (Figure 1.13). This means that the two ends of the polypeptide chain come together in the first propeller blade. It is suggested that one of the β-sheets acts as a molecular clasp, holding together and stabilising the circular structure of the propeller. The authors speculate that one side of the propeller constitutes the Ran-interaction site and the other side the DNA-interaction site (Figure 1.12 b).

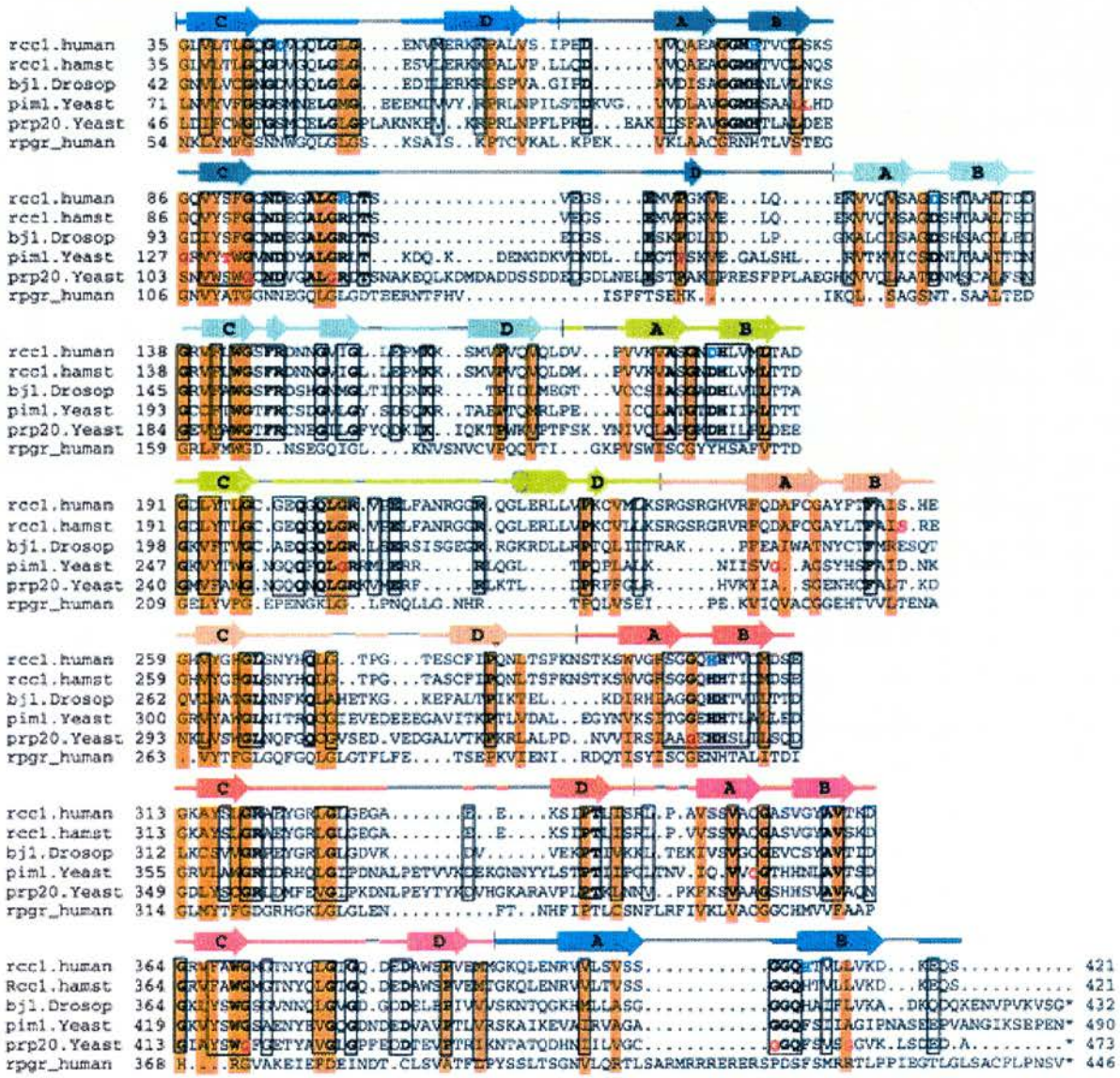


Figure 1.13: Alignment of RCC1 homologues from (from top) human, hamster, *Drosophila*, *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* with human RPGR (from Renault *et al.*, 1998). The coloured arrows indicate the residues that form the seven propeller blades shown in Figure 1.12. Residues conserved in the RCC1 homologues are boxed, residues conserved in RCC1 and RPGR are highlighted orange.

Thirty-six mutations in the *RPGR* gene were initially detected in 10-20% of XLRP patients (Meindl *et al.*, 1996; Roepman *et al.*, 1996; Miano *et al.*, 1999; Buraczynska *et al.*, 1997), well short of the 70-100% predicted by linkage studies (Ott *et al.*, 1990; Musarella *et al.*, 1990; Chen *et al.*, 1989; Teague *et al.*, 1994; Bergen *et al.*, 1995). None of these mutations were found in exons 16-19 – the majority affected conserved residues in the RCC1-homologous domain of RPGR, suggesting that this is a functionally important region. However, Vervoort *et al.* (2000) recently detected novel RPGR exons, one of which contains a mutation hotspot. They sequenced and analysed 172 kb of genomic DNA in the region of the *RPGR* gene and four new exons were identified: 15b1, 15b2, ORF14 and ORF15. The expression of these new sequences was verified with RT-PCR experiments, which also showed skipping of exons 14 and 15. Exons 15b1 and 15b2 and exon 15a, which was first identified by Kirschner *et al.* (1999), are small exons within intron 15, each of which contains a premature stop codon. Exons ORF14, which is 948 bp long, and ORF15, 2834 bp long (with a 1706 bp coding region), are derived from a 2.5 kb open reading frame situated downstream from the start of exon 14. Exon ORF14 includes the whole of exons 14 and 15 plus intron 14. Exon ORF15, which overlaps with ORF14, contains exon 15 and 2.8 kb of intron 15 sequence. A total of 17 different ORF15 mutations (representing at least 24 separate mutation events) were identified in 28 out of 47 unrelated XLRP patients. They all lead to premature termination of translation and were found within a 1061 bp stretch. None of these mutations were detected in 150 control chromosomes.

Neither the ORF14 nor ORF15 predicted amino acid sequences contain any known protein motifs. The ORF15 translation, however, contains a repetitive, purine-rich region that has been named the "plaid" domain. This domain, which is conserved in mouse, cow and *Fugu rubripes*, has unusually low sequence complexity and high glutamic acid and glycine content. It is made up of imperfect direct EEEGEGEGE repeats and has only a 2.5% pyrimidine content (Figure 1.14). The repetitive nature of the sequence, or its atypical nucleotide constitution, may account for the high mutability. For example, it is possible that this type of DNA sequence may adopt a triplex structure, a configuration that is associated with reduced replication fidelity (Wang, Seeidman, And Glazer, 1996). There are also numerous potential DNA polymerase arrest sites in exon ORF15 (Cooper and Krawczak, 1993) and since many of the ORF15 mutations involve direct repeats, pauses at these arrest sites during replication may lead to slipped-strand mispairing events (Weaver and DePamphilis, 1982).

RPGR mutations have now been found in 72% of XLRP patients, 20% in exons 1-14 and 80% in exons ORF15, accounting for at least 11% of all RP patients (Meindl *et al.*,

1996; Vervoort *et al.*, 2000). The presence of mutations in exons 1-14 and ORF15, but not in exons 16-19, in RP3 patients, suggests that a transcript containing exons 1-14 and ORF15 is the one most relevant to the disease process.

Figure 1.14: Predicted amino acid sequence of human RPGR exon ORF15, showing the repetitive plaid domain (from Vervoort *et al.*, 2000). Glutamic and/or aspartic acid (green) and glycine (red) residues are highlighted.

1.11 - Interaction of RPGR with the cGMP PDEδ subunit

Using the yeast two hybrid system, Linari *et al.* (1999) screened a mouse embryo cDNA library to identify RPGR-interacting proteins. The RPGR bait construct encoded amino acids 1-392, encompassing the whole of the RCC1-like domain. Eight out of 40 initial putative interacting clones encoded the cGMP phosphodiesterase δ subunit (PDEδ). The identities of the remaining 32 putative interacting clones were not reported. Human PDE δ was observed to interact with the human RPGR bait and with full length human RPGR (amino acids 1-815/exons 1-19, this work was carried out prior to the identification of additional RPGR exons by Vervoort *et al.* (2000)) but not with the C-terminal half of RPGR (amino acids 579-815/end of exon 14-exon 19). The interaction was confirmed by *in vitro* protein binding (pull-down) assays and affinity measurements (K_D = 90 nM). Furthermore, the presence of any of six disease-associated missense mutations in the RCC1-like domain of RPGR decreased or abolished binding to PDEδ, while a non-disease-associated missense mutation located outside the RCC1-like domain did not.

PDEδ was isolated as the fourth subunit of bovine rod PDE and is highly conserved across evolution (Hurwitz *et al.*, 1985; Figure 1.15). PDE is bound through the isoprenylated carboxy termini of the catalytic α and β subunits to rod disc membranes (Qin, Pittler, and Baehr, 1992; Anant *et al.*, 1992). The isoprenylated membrane-bound PDE is

recognised and solubilised by PDEδ under physiological conditions (Florio, Prusti, and Beavo, 1996). Since the activity of the PDE holoenzyme is not affected by this PDEδ-mediated solubilisation, it is possible that in translocating PDE from the membrane to the cytoplasm PDEδ is serving some form of regulatory role. Cook *et al.* (2000) recently reported that the addition of bovine PDEδ to a preparation of permeablised bovine rod outer segments reduced the maximal rate of light-induced cGMP hydrolysis. PDEδ has also been shown to bind a number of small GTPases, including Rab13 (Marzesco *et al.*, 1998), Rho6, Arl3 and Arl184 (Linari *et al*, 1999). It has been suggested that RPGR may act as a GEF for a small GTPase, on the basis of its homology to RCC1 and the association with PDEδ provides an indirect link with several small GTPases, although Linari *et al.* (1999) were unable to detect GEF activity in RPGR in the presence of Arl3 or Arl184. They similarly did not report a physical association between RPGR and any small GTPases.

```
human_PDEd   MSAKDERAREILRGFKLNWMNLRDAETGKILWQGTEDLSVPGVEHEARVPKKILKCKAVS
  dog_PDEd   MSAKDERAREILKGFKLNWMNLRDAETGKILWQGTEDLSVPGVEHEARVPKKILKCKAVS
bovine_PDEd  MSAKDERAREILRGFKLNWMNLRDAETGKILWQGTEDLSVPGVEHEARVPKKILKCKAVS
mouse_PDEd   MSAKDERARDILRGFKLNWMNLRDAETGKILWQGTEDLSVPGVEHEARVPKKILKCKAVS
             *********:**:**************************************************

human_PDEd   RELNFSSTEQMEKFRLEQKVYFKGQCLEEWFFEFGFVIPNSTNTWQSLIEAAPESQMMPA
  dog_PDEd   RELNFSSAEQMEKFRLEQKVYFKGQCLEEWFFEFGFVIPNSTNTWQSLIEAAPESQMMPA
bovine_PDEd  RELNFSSAEQMEKFRLEQKVYFKGQCLEEWFFEFGFVIPNSTNTWQSLIEAAPESQMMPA
mouse_PDEd   RELNFSSAEQMEKFRLEQKVYFKGQCLEEWFFEFGFVIPNSTNTWQSLIEAAPESQMMPA
             *******:****************************************************

human_PDEd   SVLTGNVIIETKFFDDDLLVSTSRVRLFYV
  dog_PDEd   SVLTGNVIIETKFFDDDLLVSTSRVRLFYV
bovine_PDEd  SVLTGNVIIETKFFDDDLLVSTSRVRLFYV
mouse_PDEd   SVLTGNVIIETKFFDDDLLVSTSKVRLFYV
             ************************:******
```

Figure 1.15: CLUSTAL W alignment (Thompson *et al.*, 1994) of the human, dog, bovine and mouse cGMP PDE delta subunit proteins (* = single, fully conserved residue, : = conservation of strong groups, . = conservation of weak groups, blank space = no consensus).

Two other groups have isolated a novel RPGR-interacting protein, identical to that described in this thesis and in Boylan and Wright (2000). The findings of Roepman *et al.* (2000) and Boylan and Wright (2000) were published at the same time and those of Hong *et al.* (2000a) subsequently. These will be discussed in chapters 3, 4 and 5.

1.12 - Mouse RPGR

Yan *et al.* (1998) isolated and characterised the murine *RPGR* homologue (*mRPGR*). The gene was found to have two in-frame initiation codons and a number of splicing variants

were detected. Northern blot studies of mouse RNA revealed a 2.7 kb transcript widely expressed in murine tissues, with greater abundance in brain and testis. A faint brain-specific band of 2.8 kb was observed, and two transcripts, one of 3 kb and one of 3.3 kb, were testis-specific. In transfected COS cells, mRPGR was found to be isoprenylated and localised to the Golgi complex.

Hong *et al.* (2000b) prepared an RPGR-deficient mouse model of RP3. Immunoblots from wild-type and knockout mice were probed with anti-RPGR antibodies and, in the wild-type only, a band of 95 kDa was detected in a variety of tissues, and an additional larger band of 165 kDa was seen in testis only. Tissue sections from the RPGR-deficient mice were stained to reveal the presence of the β-galactosidase reporter, the transcription of which was driven by the endogenous *mRPGR* promoter. In this way, evidence was found of widespread *mRPGR* promoter activity in the brain, and in the ganglion and photoreceptor cell layers of the retina. Wild-type retinas were then examined immunocytochemically to establish the normal subcellular position of RPGR. There was prominent staining in the region of the photoreceptor connecting cilium at the junction between the inner and outer segments. To investigate whether RPGR has a similar location in cones, attention was switched from mouse to squirrel. Mouse retinas contain large numbers of rods but very few cones, whereas in the squirrel, cones dominate. The immunocytochemistry data indicated that, in the squirrel retina, RPGR localises to the connecting cilium of cones. The subcellular localisation of RPGR was further investigated by examining mechanically-dissociated mouse photoreceptors. These predominantly consisted of outer segments attached to the connecting cilia, with the occasional inner segment fragment. Staining with anti-RPGR antibodies confirmed the ciliary location. This staining was absent in preparations from RPGR-deficient mice. It should be noted that the antibodies used by Hong *et al.* (2000b) recognise the carboxyl-terminal 250 amino acids of mRPGR, corresponding to the hRPGR residues encoded by exons 12 to 13 and 16 to 19 (although the nucleotide sequence of *mRPGR* is subtantially different from human *RPGR* in exons 11-17 (Yan *et al.*, 1998)). Kirschner *et al.* (1999) had previously reported that the mouse gene is subject to extensive alternative splicing, particularly the 3' half of the gene. The restricted localisation of mRPGR reported by Hong *et al.* may therefore reflect this alternative splicing. At an early age (approximately 30 days postnatal) retinal function appeared normal in the knockout mice, as measured by ERG, although the authors noted slight immunohistochemistry irregularities in rods and cones. In cones, the blue and green opsins, which are normally found in the outer segments, localised to the inner segments, perinuclear and synaptic regions. In rods, the rhodopsin content was reduced. Retinal

degenerative changes had begun by 2 months of age, and after 6 months there was substantial photoreceptor loss with reduced outer segment length and reduced photoreceptor nuclear layer thickness. In addition, there was disorganisation of the newly formed disc membranes at outer segment bases. The cilia of the knockout mice did not appear to be affected structurally.

1.13 - Aim of this thesis

The aim of this thesis was to contribute to the characterisation of the RPGR protein, mutations in which are responsible for RP3-type XLRP by identifying proteins that interact with RPGR. It was hoped that this could provide a step in the assembly of functional information about this protein and the biochemical pathways it is involved in. The initial goal was to use the RCC1-like domain of RPGR (which was affected by the majority of the RP3 mutations reported at the time this work began) in a yeast two-hybrid screen of candidate interactors and expression libraries in an effort to isolate potential interacting partners. The work that was carried out to achieve this is described in Chapter 3 (The yeast two-hybrid system, page 84). A novel RPGR-interacting protein was identified and was characterised in terms of expression, genomic structure and chromosomal localisation. The results that were obtained and the preliminary screening of this gene for mutations in a subgroup of Sardinian ARRP patients are presented in Chapter 4 (Characterisation of RPGRIP, page 116). The findings of this thesis are discussed in Chapter 5 (Discussion, page 161).

CHAPTER TWO

MATERIALS AND METHODS

2.1 - Media, antibiotics and additives.

All media were sterilised by autoclaving at 121°C/15 pounds per square inch for twenty minutes (all autoclaving was carried out in this way unless unless otherwise specified). All chemicals in this and the following sections were purchased from Sigma Chemical Co., St Louis, USA.

LB Medium/litre

Bacto-tryptone (Difco, West Molesey, UK)	10 g
Bacto yeast extract (Difco)	5 g
NaCl	10 g
Distilled water (dH_2O)	to 1 litre

Adjust pH to 7.0. For agar plates, 15 g/litre agar (Difco) was added.

M9 Minimal Medium/litre

5x M9 salts	200 ml
dH_2O	to litre
1 M $MgSO_4.7\ H_2O$	1 ml
20% glucose	10 ml
1 M $CaCl_2$	0.1 ml

For agar plates, 15 g/litre agar (Difco) was added.

5x M9 salts/litre

$Na_2HPO4.7H_2O$	64 g
KH_2PO_4	15 g
NaCl	2.5 g
NH_4Cl	5 g

Ampicillin

Ampicillin is kept at -20°C as a stock solution at 50 mg/ml in dH_2O and added to a final concentration of 50 µg/ml. Added to media following sterilisation, immediately before use, or when the temperature of the media was below 45°C.

IPTG (Isopropylthio-β-D-galactoside)

Stored at -20 °C at 200 mg/ml in dH_2O. 4 µl was spread onto an L-agar plate as required for α-complementation/blue-white colour selection (Melford laboratories Ltd, Ipswich, UK).

X-gal (5-Bromo-4-chloro-3-Indoyl-β-D-galactoside)

Stored at -20 °C at 20 mg/ml in dimethylformamide. 40 µl was spread onto an L-agar plate as required for α-complementation/blue-white colour selection (Melford laboratories Ltd).

2.2 - Bacterial Strains

DH5α (Gibco BRL, Paisley, UK) [Genotype: *supE*44, Δ(lacZYA-*arg*F)U169, ϕ80d*lacZ*ΔM15, *hsdR*17, *recA*1, *endA*1, *gyrA*96, *thi*-1, *rel*A1.] A recombination-deficient suppressing strain used for plating and growth of plasmids and cosmids.

HB101 (Gibco BRL) [Genotype: *supE*44, *hsdS*20, *recA*13, *ara*-14, *proA*2, *lacY*1, *galK*2, *rpsL*20, *xyl*-5, *mtl*-1.] A suppressing strain commonly used for large-scale production of plasmids. It is an *E. Coli* K12 x *E. Coli* B hybrid that is highly transformable.

DH10B (Gibco BRL) [Genotype: F^-, *mrc*A, Δ(*mrr-hsd*RMS-*mrc*BC), ϕ80d*lacZ*ΔM15, Δ*lac*X74, *endA*1, *recA*1, *deo*R, Δ (*ara, leu*)7697, *ara*D139, *gal*U, *gal*K, *nup*G, *rps*L, λ–.] An *E. coli* strain (Electro-competent cells) used here for amplifying a cDNA library because of the high level of transformation efficiency possible.

INVαF' One Shot (Invitrogen Corp., Abingdon, UK) [Genotype: F', *endA*1, *recA*1, *hsdR*17, *supE*44, *thi*-1, *gyr*A96, *rel*A1, ϕ80d *lacZ*ΔM15, Δ(lacZYA-*arg*F)U169.] Cells used with the Invitrogen 'TA' cloning system. Allows stable replication of high-copy number plasmids.

Key

ara: confers an inability to utilise arabinose

***deo*R**: allows uptake of large plasmids

***end*A**: abolishes non-specific endonuclease I activity and improves the quality of plasmid DNA isolations

F': F' episome, male *E. coli* host, necessary for M13 infection

***gal*K**: inability to use galactose; galactokinase mutation

***gal*U**: inability to utilise galactose; glucose-1-phosphate uridyltransferase mutation

***gyr*A**: mutation confers resistance to nalidixic acid; DNA gyrase mutation

***hsd*R**: restriction minus, modification positive – transformed DNA will not be cleaved by endogenous restriction endonucleases

***hsd*S**: restriction minus, modification minus – transformed DNA will not be cleaved by endogenous restriction endonucleases; protective methylation abolished

***lac*Y**: inability to use lactose; galactoside permease mutation

***lacZ*ΔM15**: partial deletion of β-D-galactosidase; allows α-complementation of β-galactosidase activity

***leu*B**: requires leucine for growth on minimal media; β-isopropyl malate dehydrogenase mutation

***mrc*A**: blocks restriction of methyl-cytosine-specific DNA sequences

mrr: absence of the *mrr* gene product allows more efficient cloning of DNA containing methyladenine residues

mtl: inability to use mannitol

***nup*G**: alters nucleoside uptake

***pro*AB**: requires proline for growth on minimal media

***rec*A**: prevents recombination between introduced DNA and host DNA

***rel*A**: allows RNA synthesis in the absence of protein synthesis

***rps*L**: resistance to streptomycin; mutation in protein S12 of 30S ribosomal subunit

***sup*E**: mutation suppresses amber (UAG) mutations

***thi*-1**: strains cannot make thiamine and therefore require thiamine to grow on minimal media

***xyl*-5**: confers an inability to utilise xylose

2.3 - Cloning vectors

pAS1 (Harper *et al.*, 1993): has amino acids 1-147 of the GAL4 transcriptional activator, high-level fusion protein expression driven by the *ADH1* promoter and terminated at the *ADH1* terminator. High level of expression during logarithmic growth of the yeast host cells, but then is repressed in late log phase by the ethanol that accumulates in the medium as a by-product of yeast metabolism. Fusion proteins have the haemagglutinin tag between

the cloned portion and the GAL4 domain. Expresses the *TRP1* gene in yeast, which encodes tryptophan prototrophy. Ampicillin resistant in *E. coli.* 8.5 kb in length. Carries the CYH^S2 gene, which confers sensitivity to cycloheximide.

pACTII (Durfee *et al.*, 1993): has amino acids 768-881 of the GAL4 transcriptional activator, high-level fusion protein expression driven by the *ADH1* promoter and terminated at the *ADH1* terminator. Fusion proteins have the haemagglutinin tag inbetween the cloned portion and the GAL4 domain. Expresses the *LEU2* gene in yeast, which encodes leucine prototrophy. Ampicillin resistant in *E. coli.* 7.55 kb in length.

pCR2.1 TOPO (Invitrogen): has 3'-T overhangs for ligation of *Taq*-amplified PCR products. Ampicillin and kanamycin resistance genes. Blue/white colony screening. 3.9 kb in length.

2.4 - Small scale preparation of plasmid DNA (miniprep)

To prepare small amounts of plasmid DNA (1-10 µg) from bacterial cultures the Promega Wizard miniprep kit (Promega, Southampton, UK) was used, based on the alkaline lysis method of Birnboim and Doly (1979). Generally an overnight culture was set up by inoculating 10 ml LB-amp with a single isolated bacterial colony from an agar plate and incubating at 37 °C with shaking. The cells were pelleted by spinning in a Sorvall RT 6000D centrifuge at 3000 rpm for ten minutes and the supernatant poured away. The pellet was then treated as directed by the kit protocol. Briefly:

1. The pellet was resuspended in 400 µl 'cell resuspension solution' (50 mM Tris-HCl pH 7.5, 10 mM Na_2EDTA, 100 µg/ml RNase A).
2. 400 µl 'cell lysis solution' (0.2 M NaOH, 1% SDS) added and gently mixed by inversion.
3. 400 µl 'neutralising solution' (1.32 M potassium acetate) added and mixed by inversion.
4. Tubes spun in an Eppendorf centrifuge for 5 minutes at room temperature to pellet cell debris and bacterial DNA.
5. Supernatant transferred to a syringe which had been attached to a Wizard minicolumn and 1 ml of the provided silca gel resin added.
6. Syringe plunger inserted and slurry pushed through the column.

7. Column washed with 2 ml 'column wash solution' (0.1 M NaCl, 10 mM Tris-HCl pH 7.5, 2.5 mM Na_2EDTA, 50% (v/v) ethanol).
8. Column transferred to a fresh 1.5 ml eppendorf tube and spun for 2 minutes to dry the resin.
9. Column transferred to a fresh 1.5 ml eppendorf tube and 50 µl water or Tris-EDTA (TE, 0.01 M Tris-HCl, 1mM EDTA, pH 7.5) added. Spun for 1 minute to elute the DNA.

2.5 - Large scale preparation of plasmid DNA (maxiprep)

To obtain milligram amounts of plasmid DNA, a procedure based on the methods of Birnboim and Doly (1979) and Ish-Horowicz and Burke (1981) was used. A single isolated bacterial colony was used to innoculate 30 ml of LB containing 50 µg/ml ampicillin. This was incubated at 37 °C with shaking for 8 hours (when the culture reaches late log phase, corresponding to an OD_{600} of 0.6). 25 ml of the culture was used to inoculate 500 ml of LB-amp (prewarmed to 37 °C), which was then incubated at 37 °C with vigorous shaking overnight. The bacterial cells were harvested by centrifugation at 4000 rpm for 15 minutes at 4°C and the supernatant discarded. The pelleted cells were washed in 18 ml of a sterile solution containing 50 mM glucose, 25 mM Tris-HCl (pH 8.0) and 10 mM EDTA. 2 ml of freshly prepared lysozyme (10 mg/ml; Sigma) was added and then 40 ml of a freshly prepared solution containing 0.2N NaOH and 1% SDS. The mixture was agitated and left to stand at room temperature for 5 to 10 minutes while the cells lysed. To the bottle containing the lysate was added 20 ml of an ice cold solution containing 60 ml 5M potassium acetate 11.5 ml glacial acetic acid and 28.5 ml H_2O. The bottle was shaken and incubated on ice for 10 minutes and then centrifuged at 4000 rpm for 15 minutes at 4°C. The supernatant containing the plasmid DNA was filtered through 4 layers of cheesecloth and to it was added 0.6 volumes of isopropanol. The bottle was mixed well and stored for 10 minutes at room temperature. The precipitated nucleic acid was pelleted by centrifugation at 5000 rpm for 15 minutes at 4° C. After discarding the supernatant, the pellet was washed in 70% ethanol, air dried and then dissolved in 3 ml of TE (pH 8.0). The plasmid DNA was then purifed by polyethylene glycol precipitation as described below. The 3 ml solution of nucleic acid was transferred to a 15-ml Corex tube and to it was added 3 ml of ice-cold 5M lithium chloride. The contents of the tube were mixed well and centrifuged at 9000 rpm for 10 minutes at 37 °C. The supernatant was transferred to a 30 ml Corex tube and mixed with an equal volume of isopropanol. The precipitated nucleic acids were recovered by

centrifugation at 9000 rpm for 10 minutes at room temperature. The pellet was washed carefully in 70% ethanol and air dried and then dissolved in 500 µl TE (pH 8.0) containing DNAase free pancreatic RNAase (20 µg/ml; Boehringer Mannheim, Mannheim, Germany). The solution was transferred to a 1.5-ml microfuge tube and stored at room temperature for 30 minutes. 500 µl of 1.6M NaCl containing 13% (w.v) polethylene glycol (PEG 8000) was added and the tube agitated. The plasmid DNA was recovered by centrifugation at 12,000 x g for 5 minutes in a microfuge. The supernatant was removed and the pellet of DNA dissolved in 400 µl of TE (pH 8.0). The solution was extracted once with phenol:chloroform and once with chloroform. The aqueous phase was transferred to a fresh tube and to it was added 100 µl of 10M ammonium acetate (the solution was mixed) and 2 volumes of ethanol. The tube was stored at room temperature for 10 minutes and then centrifuged at 12,000 x g for 5 minutes at 4°C. The pellet was washed in 200 µl of 70% ethanol, air dried thoroughly and finally dissolved in 200 µl of TE (pH 8.0) and the OD_{260} measured.

2.6 - Phenol-chloroform extraction and ethanol precipitation of DNA

Phenol-chloroform extraction: DNA was extracted from solution (following restriction enzyme digestion, etc) once by addition of an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1). The mixture was vortexed vigorously and centrifuged at 12,000 rpm for 5 minutes The upper, aqueous phase was carefully removed, avoiding precipitated proteins found at the interface between the two layers. The DNA was then extracted in the same way with one volume of chloroform.

Ethanol precipitation: To concentrate DNA and remove salts, a 1/10 volume of 3M sodium acetate pH 5.2 was added to the DNA solution, followed by 2.5 volumes of absolute ethanol. 20 mg of glycogen (Boehringer) was added as carrier when necessary. The mixture was vortexed vigorously and left at -20°C for between 1 hour and overnight. The tube was centrifuged at 12,000 rpm for 30 minutes at 4°C. The supernatant was discarded, and the pellet washed with 70% ethanol. The tube was centrifuged at 12.000 rpm for 10 minutes at 4°C. The supernatant was discarded, and the pellet air dried. The pellet was then resuspended in the appropriate volume of TE or dH_2O.

2.7 - Agarose gel electrophoresis

Electrophoresis solutions

0.5 x TBE: 45 mM Tris-borate, 0.1 mM EDTA pH 8.0

6 x DNA Loading Buffer: 0.25% Bromophenol blue, 30% glycerol in H_2O

Size marker

2 µl (250 ng/µl) of the Boehringer Mannheim molecular weight marker X was used to compare fragment sizes on agarose gels. The sizes of the fragments (in bp) are: 12216, 11198, 10180, 9162, 8144, 7126, 6108, 5090, 4072, 3054, 2036, 1636, 1018, 517, 506, 396, 344, 298, 220, 201, 154, 134.

Agarose gel electrophoresis

DNA was separated according to size in horizontal agarose gels ('Hi-Pure' Low EEO agarose, BioGene, Kimbolton, UK) using Anachem horizontal electrophoresis units. The amount of agarose used to make the gel depended on the size range of the DNA to be resolved but most commonly a 0.9% gel was prepared. All agarose gels were made with 0.5 x TBE. To stain the DNA, ethidium bromide was added to agarose gels to a final concentration of 1 mg/ml. One-sixth of the sample volume of 6 x Loading Buffer was added to the DNA prior to loading the sample on the gel. Gels were run at 30-150V depending on resolution and running time required, in 0.5 x TBE. DNA fragments were visualised on a UV transilluminator and photographed using a video copy processor (Mitsubishi, Hatfield, UK).

2.8 - Purification of DNA from agarose gels - Geneclean

To purify DNA from agarose gel slices the Geneclean spin kit (BIO101, Harefield, UK) was used, in which DNA became bound to a silica matrix while the molten agarose was eluted. The kit protocol was followed. Briefly:

1. The gel slice was put into the upper cavity of a Geneclean 'Spin filter' and 400 µl glassmilk added. Spin filter was floated in a 55°C water bath for 5 minutes to melt the agarose, and inverted every minute to prevent sedimentation of the glassmilk.
2. The filter was spun in an Eppendorf centrifuge for 1 minute. Flow through was discarded.

3. 500 µl of the provided Geneclean wash solution was added and the filter spun again for 1 minute. Flow through was discarded.
4. Filter was spun for 2 minutes to dry the glassmilk.
5. Between 10 and 30 µl elution solution (nuclease free water) was added to the filter, which was spun for 1 minute to elute the DNA.

2.9 - Purification of DNA from solution - Geneclean

To purify small amounts of DNA from solution the Geneclean spin kit was used. The protocol is the same as for purifying DNA from agarose gels except that the intial 55°C incubation is carried out at room temperature.

2.10 - Restriction endonuclease digestion of DNA

Restriction endonucleases were purchased from Boehringer Mannheim and New Enland Biolabs (Bishops Stortford, UK) and came provided with an optimal 10 x buffer stock. Digestion reactions were generally carried out in 20 µl volumes in 1 x buffer at the directed temperature (usually 37°C) for between 2 and 4 hours. The volume of enzyme added to the digest was no more than 10% of the total reaction volume to keep the glycerol concentration to below 5% (v/v) and prevent non-specific ("star") activity.

2.11 - Subcloning into plasmids

2.11.1 - Ligation

Ligation reactions were performed using the Boehringer Mannheim Rapid Ligation kit. Generally, equal masses of insert and vector DNA were used and the following protocol was followed:

1. Add 1 µl vector DNA and 7 µl insert DNA to a 1 ml Eppendorf microfuge tube. Add 2 µl of the provided 5x DNA dilution buffer and mix well.
2. Add 10 µl of the provided ligation buffer and mix well.
3. Add 1 µl T4 DNA ligase (5 U/µl), mix well and briefly spin down. Incubate at room temperature for 5 minutes.
4. Between 1 and 5 µl of the ligation reaction is transformed into 50 µl competent cells.

2.11.2 - Transformation of competent cells by heat shock

Competent cells were transformed according to the manufacturer's instructions. Briefly: for routine subcloning, 50 µl aliqouts of *E. coli* strain DH5α cells were incubated with 5 µl of ligation reaction products on ice for 30 minutes. The cells were transferred to a 37°C water bath for 20 seconds and then returned to ice for 2 minutes. 950 µl LB medium at room temperature was added and the cells incubated at 37 °C with shaking for 1 hour. The cells were pelleted by brief centrifugation and resuspended in approximately 50 µl supernatant. The cells were then spread onto fresh L-agar plates containing 50 µg/ml ampicillin and incubated at 37 °C overnight.

2.11.3 - Transformation of competent cells by electroporation

A 40 µl aliquot of electro-competent DH10B or HB101 cells was thawed on ice and transferred to a chilled, sterile, electroporation cuvette. The DNA, resuspended in a low ionic strength medium such as TE, was added to the cuvette, and the mixture subjected to a pulse of 2.5 kV (200 Ohms, 25 µF) in a Bio-Rad Gene Pulser (Bio-Rad Laboratories Ltd., Hemel Hempstead, UK). This produced a pulse with a time constant of 4.5 to 5 msec. 900 µl of LB medium was immediately added to the cells, which were transferred to an eppendorf tube and incubated for 1 hour at 37°C in shaking incubator. The cells were pelleted by brief centrifugation and resuspended in approximately 50 µl supernatant. They were spread onto fresh pre-warmed L-agar plates containing 50 µg/ml ampicillin and incubated at 37°C overnight.

2.11.4 - 'TA' cloning

The TOPO TA Cloning kit (Invitrogen) takes advantage of the non-template dependant terminal transferase activity of *Taq* polymerase, which adds a single deoxyadenosine to the 3' ends of PCR products. These 3-A overhangs are used to insert the PCR product into a vector (pCR2.1-TOPO) which has single 3'deoxythymidine overhangs at its insertion site. In the case of PCR products that had been generated by proof-reading enzymes, which remove the 3'-A overhangs, it was necessary to add overhangs proir to subcloning. 7 µl of gel purified PCR product was incubated with 1 µl *Taq* polymerase, 1 µl of dATP (2mM) and 1 µl 10x *Taq* polymerase buffer at 72°C for 10 minutes. 2 µl of this was incubated with 0.5

µl of pCR2.1-TOPO vector and incubated at room temperature for 5 minutes. 0.5 µl of 6x STOP solution (0.3 M NaCl, 0.06 M $MgCl_2$) was added and the solution mixed for approximately 10 seconds at room temperature. The entire reaction was then used to transform 50 µl competent DH5α cells. The transformed cells were plated onto L-agar plates containing 50 µg/ml ampicillin which had been pre-spread with 40 µl of 20 mg/ml X-gal and 40 µl of 100 mM IPTG. Blue/white colour selection was used to identify colonies containing recombinant plasmids. The positive colonies were used to inoculate 10 ml aliquots of LB containing 50 µg/ml ampicillin, which were incubated overnight at 37°C and then used to prepare plasmid DNA.

2.11.5 - PCR screening for positive colonies

In the absence of blue-white colour selection, a PCR screening technique was used to quickly identify bacterial colonies containing recombinant plasmids. This was carried out as follows: a PCR master mix was prepared (using primers specific for the plasmid insert), that was sufficient for ten reactions and was aliquoted into separate numbered tubes. Using filtered pipette tips, half of a separate colony was added to each of the ten tubes. The other half of each colony was used to inoculate one of ten numbered tubes containing 1.5 ml LB containing 50 µg/ml ampicillin. The PCR reaction was carried out and the results analysed using agarose gel electrophoresis. The 1.5 ml cultures corresponding to the positive PCR results were used to prepare glycerol stocks and plasmid DNA.

2.11.6 - Preparation of glycerol stocks of bacterial transformants

A single isolated bacterial colony from an L-agar plate was used to inoculate an overnight culture. The next day, 820 µl of the culture was mixed thoroughly with 180 µl sterile glycerol and put into storage at -70°C.

2.12 - Polymerase chain reaction

Oligonucleotides were designed to have a melting temperature (T_m) of between 55°C and 70°C, if possible, and were synthesised by Genosys (Cambridge, U.K.). The lyophilised pellets were resuspended in sterile dH_2O at a concentration of 100 µM and, from these, working stock solutions of 10 µM were prepared and stored at -20°C

For most PCR applications, the Expand High Fidelity PCR system (Boehringer

Mannheim) was used according to the manufacturer's instructions. This incorporates a mix of *Taq* and *Pwo* DNA polymerases. Reaction volumes of 25 or 50 µl were used, in 0.5 ml microfuge tubes. Reactions were carried out in either an MJ Research (Waltham, USA) PTC-200 Peltier thermal cycler (using a heated lid) or a Perkin Elmer-Cetus (Beaconsfield, UK) DNA thermal cycler (in which case reaction mixtures were overlaid with 1 drop of light mineral oil (Sigma).

PCR programmes had the following cycling parameters: (i) an initial single denaturation step of 2 minutes at 95°C, followed by (ii) 30 cycles consisting of a denaturation step of between 15 seconds and 1 minute at 95°C, an annealing step of between 30 seconds and 1 minute, if possible at T_m - 5°C and an extension step of variable duration according to the size of the product, at 72°C, then (iii) a final extension step of 7 minutes at 72°C.

The duration of the extension step varied according to the expected size of the PCR product:

Extension time	45s	1 minute	2 minutes	4 minutes	8 minutes
Size of PCR product (kb)	Up to 0.75	1.5	3	6	10

One fifth of the volume of the PCR reaction was analysed by agarose gel electrophoresis. Oligonucleotide primers used in this project are listed in Table 2.1.

Table 2.1: Oligonucleotide PCR primer information.

Name	**T_m (°C)**	**Sequence (5' to 3')**
1048 Fwd	69.8	AACATGGCAAAAGAATAGGTGTTCAAGG
1153 Fwd	66.7	CAGCAGGTGAATTACACTGAGTGG
139 Fwd	71.0	TGTCATAGTGCCACCCATGTCTCAG
19 Rev	67.3	TGTAGAGGTTCCTTCTGTTTCAGTGC
-26 Fwd	67.8	CATATCACACCCCTTGGTATCCATC
4 Rev	67.5	TGTTTCAGTGCTGAATGGGATAATG
419 Fwd	57.6	GAAGAAATCCACTTTCACTTTAG
430 Fwd	54.1	AGAGAACACTCAGATCGAG

44 Rev	65.9	TTCAGAGGATTCTTTGTCATTTACAGG
?4/5 junction REV	71.1	CAGTGTTTCCAAACATAACGACAACGG
494 Rev	55.6	GTGATACTTGCAACTCATTTAG
522 Rev	64.0	GATCAGGATCTTGTCCATTCAG
544 Rev	54.8	CTAACAATGTCTAGCTCTTGCTCTAG
-6 Fwd	63.0	TGAGAGCTTGCTTTCCATTG
638 Fwd	43.8	CTGACAATGATAATACAG
670 Rev	65.6	CTTAGATCGGTTATTCTTGAGGTCTAGATG
686 Fwd	68.8	GAGGTTGACTCGACATCTAGACCTCAAG
761 Rev	70.0	CTTCAGTCATCTCCTTGTAAATAGC
801 Fwd	63.7	TGGATTGGAAGTTTCCCTACATAC
845 Rev	70.7	GTATCCTTCCCCTTAGTCTGAGCTTCTGG
869 Fwd	72.3	GAAACCAGAAGCTCAGACTAAGGGGAAGG
962 Rev	69.3	AGGTTTTCTCTTAGATCGATACTGGCCAG
ARF1ATGNcoI	79.5	ACAACCATGGGGAACATCTTCGCCAACCTC
ARFStopBamHI	75.9	TTGGATCCTTCACTTCTGGTTCCGGAGCTG
Bov Fwd	64.6	CCCTAAGGCAGATTCAGAGAAGAT
Bov Rev	66.4	GCAGCAGCTTGAAGAGACACC
BovG3PDHAS	62.0	ATCGCTGTTGAAGTCGCAGG
BovG3PDHS	56.0	TGAACCACGAGAAGTATAAC
Exon12BamHI	59.1	CGGGATCCTACATGTTTAAGATATCAGTAGTT
Exon13NcoI	86.0	TAATCCATGGATACACACATCATGAGCCTG
G3PDHF	66.0	GACCACAGTCCATGACATCACT
G3PDHR	66.0	TCCACCACCCTGTTGCTGTAG
Genscan exon1	60.5	CATGTATTACTTTATTCGGATCAAAG
Genscan exon10	64.6	TGAAAGAAATGCTTCATTGGTTATG

Genscan exon11	67.0	GAAGAGCCAACTGGAAGATGTGTC
Genscan exon12	61.1	GAAAATTGCTGAATGACAATTATG
Genscan exon17	67.5	GCTCAAAGATGTTGCTTATGGCAC
Genscan exon18	66.9	ACAGGCGTGCAATAAACGAAAG
Genscan exon19	76.1	CCCAGTGCTTGTGACCTCTGACCTGG
Genscan exon2	58.9	GAGATTTAATTACAAACGTAAGCTG
Genscan exon3	67.4	GAGCATCACGATCTTCATTTTCATG
Genscan exon4	79.9	CTGCCGCCGAGGCGAAGGCCTAG
Genscan exon5	71.3	CACATCTGGTGGACCCTACATCAGG
Genscan exon6	68.1	ACATGTTGGTGAAGGAGCTTTCTTG
Genscan exon7	80.7	TGCAAGGGCATTTCCACTGCGTCGG
Genscan exon8	66.7	GAGGTGAAGTAGCCAGTAAACCCAG
Genscan exon9	70.1	GCTCATTGGAGTGTGCTCAGAAGG
Intron 3 Fwd	74.6	GCAGTCTTCTTGACCTAGCCAGTGCCAC
Intron 4 Fwd	69.0	AAAATGGGAAGGAAGTAGATAAGGTGCTG
Intron 4 Rev	67.0	GATAAGTCATATGCTAAGGGTATGAGAGGC
Intron 5 Fwd	78.0	GGAGCGAAAGCCTGGGTTTTACTGCGG
Intron 5 Rev	82.4	GCTGGGATTACAGGCGTGAGCCACTGCACC
M13F	64.0	CCCAGTCACGACGTTGTAAAACG
M13R	58.0	AGCGGATAACAATTTCACACAGG
Intron 6 Rev	85.7	TACAGGCCGAAGCCACCATGCCCGGCC
pACTinsFwd	68.9	CCCATACGATGTTCCAGATTACGC
pACTinsRev3	65.2	ATGCACAGTTGAAGTGAACTTGC
PDEdBamHI	69.7	CAAGGATCCAACATAGAAAAGTCTCACTCTG
PDEdNcoI	76.2	ATATCCATGGTCATGTCAGCCAAGGACGAG
pGAD10Fwd	57.6	ATCTATTCGATGATGAAGATACC

pGAD10Rev	62.5	GCACAGTTGAAGTGAACTTGC
PT3T7DF	69.8	CTATAGGGAATTTGGCCCTCGAGG
PT3T7DR	62.2	TAACCCTCACTAAAGGGAATAAGC
r93221 Fwd 2	64.1	GAACAAGGTTCTGAAGTCAGTGAAG
r93221BamHI	74.0	AAGGATCCATGTTCTCATCAGACATC
r93221NcoI	74.0	TATACCATGGTTATGACATTATCCCATTCAGCA
Rab13BamHI	80.0	ATGGATCCTCAGCCCAGGGAGCACTTGTTG
Rab13NcoI	77.1	TATGCCATGGCCAAAGCCTACGACCAC
RanNcoI	74.4	ATAACCATGGCTGCGCAGGGAGAG
RanStopBamHI	73.6	CATGGATCCTTCACAGGTCATCATCCTC
RCC1ATGNcoI	82.1	AAGGCCATGGAGATGTCACCCAAGCGCATAGC
RCC1StopBamHI	75.5	AGGGATCCTCAGCTCTGTTCTTTGTCCTTGAC
Rev 3	72.0	TTCAATGCACATCTTCTCTGAATCTG
RG32 Bam HI	88.0	GAGGATCCCAGAGACTTTTAGATTGGAATAG
RG32 long Rev	71.5	GCCATCACCTATGAAGCTGTTAGTCTCTGAG
RPGRATGNcoI	77.3	TAATCCATGGTAATGAGGGAGCCGGAAGAGC
Stop +36	57.3	TTTACTATGGTTCCCTCAGAG
Stop +6:	54.4	TTAGATTGGAATAGCACTTG
StopBamHI	73.6	CATGGATCCTTCACAGGTCATCATCCTC

2.13 - Reverse Transcriptase PCR

The Access RT-PCR kit (Promega) was used to amplify *RPGRIP* from samples of total human or bovine RNA using gene specific primers as indicated in Chapters 3 and 4. The kit uses AMV reverse transcriptase from Avian Myeloblastosis Virus for first strand DNA synthesis and *Tfl* DNA polymerase for second strand cDNA synthesis and DNA amplification. 25 µl reactions were set up according to the manufacturer's instructions. RNA samples were provided by Dr. Raf Vervoort and Sandy Edgar (both MRC Human

Genetics Unit, Edinburgh). The reactions were carried out in an MJ Research PTC-200 Peltier thermal cycler using the following programme: 48 °C for 45 minutes; 94 °C for 2 minutes; 35 cycles of 94 °C for 1 minute, 60 °C (for *RPGRIP*)/55°C (for *G3PDH*) for 1 minute, 72 °C for 1 minute; then 72 °C for 10 minutes.

The primers used were:

Product	Forward primer	Reverse primer
Human *RPGRIP*	1048F	544Rev
Human *G3PDH*	G3PDHF	G3PDHR
Bovine *RPGRIP*	BOV FWD	BOV REV
Bovine *G3PDH*	BovG3PDHS	BovG3PDHAS

2.14 - Sequencing of DNA

2.14.1 - Method 1

DNA was sequenced using the BigDye terminator cycle sequencing ready reaction kit (PE Applied Biosystems, Beaconsfield, UK). Typically, a reaction was made up of 4 µl of the provided reaction mix, 4 µl 2.5x buffer (200 mM Tris-HCl pH 9.0, 5 mM $MgCl_2$), 15 ng (PCR product) or 1 mg (plasmid) DNA template, 3.2 pmol primer and dH_2O to 20 µl.

Prior to sequencing plasmid DNA, it was first linearised by enzymatic digestion, phenol extracted and ethanol precipitated. If a PCR product was sequenced, the DNA was purified using the Geneclean spin kit. 30 cycle sequencing reactions were carried out in a Perkin Elmer-Cetus DNA thermal cycler using the following program: 94°C for 30 seconds, 50°C for 15 seconds and 60°C for 4 minutes. The samples were ethanol precipitated and the pellet washed with 70% ethanol and air dried. The samples were loaded by Agnes Gallacher (MRC Human Genetics Unit) on an ABI 377 automated sequencer, and the results analysed using the ABI Sequencing Analysis 3.0 software.

2.14.2 - Method 2

The Thermo Sequenase radiolabelled terminator cycle sequencing kit (Amersham Pharmacia Biotech, Little Chalfont, UK) was also used to sequence DNA. To prepare the termination master mixes (G,A,T and C) 2 µl of the provided dGTP nucleotide master mix was mixed on

ice with 0.5 µl of the [α–^{33}P]dideoxynucleotide (ddNTP) terminators as shown in the table below.

To prepare termination mixes for (n) reactions:

	G	A	T	C
Nucleotide master mix	(2 x n) µl	(2 x n) µl	(2 x n) µl	(2 x n) µl
[α–^{33}P]ddNTP	(0.5 x n) µl	(0.5 x n) µl	(0.5 x n) µl	(0.5 x n) µl
Total	(2.5 x n) µl	(2.5 x n) µl	(2.5 x n) µl	(2.5 x n) µl

Four tubes were labelled G,A,T or C and filled with 2.5 µl of the respective termination mix. The individual reaction mixtures were prepared by mixing 2 µl of the provided reaction buffer, 50-500 ng template DNA, 0.5-2.5 pmol primer, 2 µl thermo sequenase polymerase (4 U/µl) and H_2O to 20 µl. The cycling termination reactions were set up by adding 4.5 µl of the reaction mixture to each of the 4 tubes containing the termination mixes. The contents of the tubes were mixed well and overlayed with one drop of light mineral oil (Sigma). The 30-cycle reactions were carried out in a Perkin Elmer-Cetus DNA thermal cycler using the following program: 95°C for 30 seconds, 55°C for 30 seconds, 72°C for 1 minute. 4 µl of stop solution (containing 95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF) was added to each tube, which was then vortexed and centrifuged briefly to separate the mineral oil from the aqueous phase. Before loading 3-5 µl of each reaction onto the gel the samples were heated to 72°C for 2-10 minutes. Gels were run at 50 W for between 1.5 hours and 5 hours depending on which region of the sequence needed to be read.

Pre-treatment of PCR templates

Before sequencing PCR products they were treated with shrimp alkaline phosphatase (Amersham Pharmacia Biotech) and exonuclease I (Amersham Pharmacia Biotech). 5 µl out of the 50 µl PCR product was mixed with 1 µl of shrimp alkaline phosphatase and 1 µl exonuclease I and incubated at 37°C for 15 minutes and then at 80°C for 15 minutes.

2.14.3 - Preparation of electrophoresis gels

Solutions

8% acrylamide:

40 g acrylamide/bis-acrylamide (19:1) (BDH Ltd. Poole, UK)

250 g urea (Gibco BRL)

25 ml 20 x glycerol tolerant gel buffer

H_2O to 500 ml

20 x glycerol tolerant gel buffer:

108 g Tris base

36 g taurine (ICN Biomedicals, Inc., Basingstoke, UK)

2 g $Na_2EDTA.2H_2O$ (BDH Ltd)

H_2O to 500 ml.

A Bio-Rad sequi-gen sequencing cell apparatus was used to prepare electrophoresis gels. Prior to assembly of the apparatus, 2-5 ml of dimethyldichlorosilane solution was applied to the fixed glass plate to encourage the gel to preferentially adhere to the free glass plate when dismantling the apparatus after electrophoresis. To plug the bottom gap between the plates the apparatus was placed in a casting tray and a mix of 15 ml 8% acrylamide, 240 µl ammonium persulphate and 40 µl N,N,N',N'-Tetramethylethylene-diamine (TEMED) (Gibco BRL) was poured in and allowed to set for 20 minutes. The gel was poured (containing 45 ml 8% acrylamide, 350 µl ammonium persulphate and 50 µl TEMED) and allowed to set for 1.5 hours.

2.15 - Autoradiography

Autoradiographs were prepared by exposing radioactive gels to Kodak Biomax film (Kodak, Cambridge, UK) overnight or for longer periods, and then processing the film in a Fuji RGII X-ray processor (Fujifilm, Bedford, UK).

2.16 - Northern Blotting

When preparing or probing northern blots, precautions were taken to prevent ribonuclease (RNase) contamination. Disposable gloves were always worn and were changed frequently, sterile, disposable plasticware was used, and solutions were treated by adding diethylpyrocarbonate (DEPC) to 0.01%, leaving overnight and then autoclaving.

2.17 - Preparation of RNA

Samples of human tissue from a recently deceased 83 yr old female were obtained from the Western General Hospital, Edinburgh. 100 mg pieces were removed from each sample and used to prepare total RNA using TRIzol Reagent (Gibco BRL), a mono-phasic solution of phenol and guanidine isothiocyanate. Human retina tissue was obtained separately by Dr. Julie Kenyon (MRC Human Genetics Unit, Edinburgh) and then treated in the same way as the other tissues. The procedures were carried out at room temperature unless otherwise stated. The pieces of tissue were each homogenised in 1 ml of TRIzol Reagent in a glass homogeniser, then removed to sterile microfuge tubes and allowed to stand for 15 minutes. 0.2 ml chloroform was added to each tube, which was then shaken vigorously for 15 minutes and left for 2 to 3 minutes before being centrifuged at 12,000 x g for 15 minutes at 4°C. The aqueous phase was transferred to a fresh tube and the RNA precipitated by adding 0.5 ml isopropanol. After 10 minutes the tubes were centrifuged at 12,000 x g for 10 minutes at 4°C. The supernatant was removed and the pellet of RNA washed once with 1 ml of 75% ethanol. The tubes were vortexed and centrifuged at 7,500 x g for 5 minutes at 4°C. The pellet was briefly air-dried and then redissolved in DEPC-treated dH_2O. The RNA yield was estimated by measuring the absorbance at 260 nm and the purity of the preparations assessed by measuring the A_{260}/A_{280} ratio.

2.18 - Poly(A)$^+$ RNA purification

Poly(A)$^+$ RNA was purified from the total RNA preparations by affinity purification on oligo(dT)-cellulose using the Pharmacia Biotech mRNA purification kit. Oligo(dT)-cellulose columns were inverted several times and the top and bottom closures removed. The columns were placed in a 50 ml centrifuge tube and the storage buffer allowed to drain away. 1 ml of the provided high-salt buffer (10 mM Tris-HCl pH 7.5,1 mM EDTA, 0.5 M

NaCl) was added and allowed to drain through under gravity. This step was repeated and then the centrifuge tube emptied of the collected buffers. The RNA samples were heated at 65°C for 5 minutes, then placed on ice and supplemented with 0.2 ml of the provided sample buffer (10 mM Tris-HCl pH 7.5,1 mM EDTA, 3.0 M NaCl). The samples were applied to the top of the cellulose beds in the columns and allowed to soak in under gravity. The columns were centrifuged at 350 x g for 2 minutes. 0.25 ml of the high-salt buffer was added and the columns spun again. This high-salt wash step was repeated. The column was then washed with 0.25 ml of low-salt buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.1 M NaCl) and centrifuged at 350 x g for 2 minutes a total of three times. The collected buffers were removed from the centrifuge tubes and replaced with sterile 1.5 ml microcentrifuge tubes. The poly(A)$^+$ RNA was eluted from the columns by washing with four successive 0.25 ml aliquots of the provided elution buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA), prewarmed to 65°C. The tubes were centrifuged at 350 x g for 2 minutes after each application and the entire eluate collected in the same tube. According to the manufacturer, after one such round of purification, approximately 50% of the recovered RNA is poly(A)$^+$ RNA. Again, the RNA yield was estimated by measuring the absorbance at 260 nm and the purity of the preparations assessed by measuring the A_{260}/A_{280} ratio.

2.19 - Electrophoresis of RNA through agarose gels

The method used to prepare northern blots was based on the protocols of Sambrook *et al.* (1989).

Solutions

5 x Formaldehyde gel running buffer

0.1 M MOPS (pH 7.0)

40 mM sodium acetate

5 mM EDTA (pH 8.0)

Formaldehyde gel-loading buffer

50% glycerol

1 mM EDTA (pH 8.0)

0.25% bromophenol blue

0.25% xylene cyanol FF

A 1.4% gel was poured using agarose reserved exclusively for RNA work. 1.4 g of agarose ('Hi-Pure' Low EEO agarose, BioGene) was melted in 62 ml of DEPC-treated H_2O. It was then cooled to 60°C before the addition of 20 ml 5 x formaldehyde gel running buffer and 18 ml formaldehyde (Fisher Scientific, Manchester, UK). The 6 µg RNA samples (containing approximately 50% poly(A)$^+$ RNA) were prepared by mixing them with 25 µl formamide, 8.75 µl formaldehyde and 5 µl 5 x formaldehyde gel running buffer. The samples were incubated at 65°C for 15 minutes, chilled on ice and then briefly centrifuged. 5 µl of formaldehyde gel-loading buffer was added to each sample. The gel was pre-run in 1 x formaldehyde gel running buffer for 5 minutes at 75 V before loading, and then at 60 V until the bromophenol blue had migrated 10 cm.

2.20 - Northern blotting

The RNA was tranferred from the agarose gel to a Zeta probe GT nylon blotting membrane (Bio-Rad). A deep rectangular tray was filled with 10 x SSC buffer. A rectangular sheet of glass was placed on top of the tray with its long axis perpendicular to that of the tray. 3 sheets of 3MM Whatman (Maidstone, UK) paper were placed on top of the glass plate so that the ends of the paper were immersed in the buffer in the tray. The paper was flooded with buffer. The agarose gel containing the separated RNA was put on top of the paper and covered with Saran wrap. A window was cut in the Saran wrap with a sterile razor blade, allowing only the gel to emerge from the window. The surface of the gel was wetted with 10 x SSC and the membrane placed onto the gel. One corner was cut from the gel and the membrane to allow the membrane to be correctly orientated at a later time. Care was taken to remove any bubbles in the assembly. The top of the membrane was wetted with 10 x SSC and then covered with 2 pieces of 3MM Whatman paper of the same size which had been pre-soaked in 10 x SSC. A large stack of paper towels was placed over the 3MM paper and a glass plate put on top. Transfer was allowed to continue overnight. The membrane was removed and briefly washed in 2 x SSC. The RNA was fixed onto the membrane in a Spectrolinker XL-1500 ultraviolet crosslinker using the automatic crosslink setting (Spectronics Corporation, New York, USA) and the membrane air dried. The marker lane was cut off and soaked in 1 M acetic acid for 15 minutes and then immersed for 45 minutes in a staining solution containing 100 mg methylene blue, 25 ml 0.4 M sodium acetate and 25 ml 0.4 M acetic acid. It was rinsed in dH_2O to allow the bands to develop.

2.21 - Preparation of radiolabelled probes

The High Prime reagent (Boehringer Mannheim) was used to prepare radiolabelled nucleic acid probes by a random primer extension method (Feinberg and Vogelstein 1983). High Prime contains random hexameric oligonucleotides, Klenow polymerase, dATP, dGTP, dTTP and a reaction buffer. Typically 50-100 ng denatured DNA template was added to 4 µl High Prime and 5 µl [α^{32}P] dCTP (50µCi) and incubated at 37°C for 15 to 30 minutes. The reaction was stopped by adding 2 µl 0.2 M EDTA (pH 8.0). The radiolabelled probe was separated from unincorporated counts using gravity flow chromatography by passing it through a Sephadex G-50 Nick column (Pharmacia Biotech) that had been pre-equilibrated with TE and eluting in 400 µl TE. The activity of an aliquot of the purified probe was measured in a Packard Tri-Carb 1500 liquid scintillation analyser.

2.22 - Hybridisation of radiolabelled cDNA probes to northern blots

A PCR fragment consisting of the 3'-half of RPGRIP (from the middle of exon 10 to the end of exon 15) was used to generate a [α-^{32}P]dCTP-labelled radioactive probe. A human multiple tissue northern blot (MTN2; Clontech, Basingstoke, UK) was incubated with 10 ml Ultrahyb hybridisation solution (containing 50% formamide; Ambion, Austin, USA) for 1 hour at 42 °C. The *RPGRIP* 3' probe was added to the hybridisation solution and incubated overnight at the same temperature. The blot was then washed twice for five minutes each in 2 x SSC, 0.1% SDS at 42°C and then exposed to a phosphor screen overnight. The screen was processed using the Molecular Dynamics Storm 840 imaging system (Amersham Pharmacia Biotech).

A second radioactive probe was prepared using a PCR fragment comprising exons 1-4 and 8-15 of *RPGRIP* as the template. This was hybridised to another human multiple tissue northern blot (MTN; Clontech), which had been incubated in 5 ml Expresshyb, a premade aqueous hybridisation solution (Clontech), for 30 minutes at 68°C. The blot was incubated with the probe at the same temperature for one hour. The blot was then washed three times, for ten minutes each, in 2 x SSC, 0.05% SDS at room temperature and exposed to a phosphor screen overnight. The screen was processed as above.

The two commercial northern blots were then hybridised to a control [α-^{32}P]dCTP-labelled radioactive human *G3PDH* probe (obtained by radiolabelling a PCR product generated from a retinal cDNA library (Clontech, Basingstoke) using the G3PDHF and G3PDHR oligonucleotides, table 2.1). The blots were prehybridised in 10 ml Expresshyb

for 30 minutes at 68°C. The probe was added and the blots incubated for an hour at 68°C. The blots were then washed three times, for ten minutes each, in 2 x SSC, 0.05% SDS at room temperature and exposed to a phosphor screen overnight. The screen was processed as before.

A *RPGRIP* PCR product (from the FWD2 and 761 oligonucleotide primers) consisting of exons 11 to 15 was used to prepare a [α-^{32}P]dCTP-labelled probe. This probe was hybridised to a northern blot containing poly(A)$^+$ RNA purified from a range of human tissues. The blot was incubated in 15 ml Ultrahyb hybridisation solution for 1 hour at 42 °C and then the probe was added and the blot left at the same temperature overnight. The blot was washed twice for five minutes in 2 x SSC, 0.1% SDS and then once in in 0.1 x SSC, 0.1% SDS at room temperature. The blot was exposed to Kodak Biomax X-ray film for 4 days. This blot was then probed with radiolabelled β-actin. A random-primed [α-^{32}P]dCTP-labelled probe was prepared from a commercial actin template (Clontech). The blot was pre-hybridised in 15 ml Ultrahyb hybridisation solution for 1 hour at 42 °C. The probe was added to the solution and left overnight at 42 °C. The probe was washed twice for 5 minutes in 2 x SSC, 0.1% SDS and then twice for 15 minutes in 0.1 x SSC, 0.1% SDS, at 42 °C. The blot was exposed to a phosphor screen for 30 minutes.

2.23 - Tissue *in situ* hybridisation

Solutions:

Hybridisation buffer	(1 ml)
10 x salt	100 µl
Ultrapure formamide	500 µl
50% dextran sulphate	200 µl
yeast RNA (10 mg/ml)	100 µl
50 x Denhardt's solution	20 µl
dH_2O	80 µl

10 x salt (1 litre)

NaCl	114 g
Tris-HCl	14.04 g
Tris Base	1.34 g
$NaH_2PO_4.H_2O$	7.8 g

Na_2HPO_4	7.1 g
0.5 M EDTA	100 ml

Make up to 1 litre with dH_2O, pH should be approximately 7.5, autoclave.

Solution A

SSC	1 x
formamide	50%
Triton X-100/Tween 20	0.1%

10 x TBST

1.4 M NaCl

27 mM KCl

0.25 M Tris-HCl pH 7.5

1% Triton X-100/Tween 20

NTMT

100 mM NaCl

100 mM Tris-HCl pH 9.5

50 mM $MgCl_2$

0.1% Triton X-100/Tween 20

Nitro blue tetrazolium (NBT)

75 mg/ml in dimethylformamide (store at -20°C)

5'-Bromo-4-chloro-3-indolylphosphate (BCIP)

50 mg/ml in 70% dimethylformamide (store at -20°C)

A mouse IMAGE EST clone was obtained (accession number aa204546) was obtained from the HGMP RC. The plasmid backbone, pT3T7-Pac, has the T7 promoter immediately upstream of the polylinker and the T3 promoter downstream. The orientation of the clone was checked by PCR-amplifying between the 544 REV primer (this primer will anneal to the sense strand of the mouse *RPGRIP* cDNA) and either the pT3T7-FWD or pT3T7-REV primers (these primers anneal to the T7 and T3 promoters respectively). If the insert was in the correct orientation only the 544 REV/pT3T7-FWD reaction would give a product. This

was found to be the case. Two aliquots of aa204546 DNA were seperately cut with either *Hind*III or *Xho*I and then phenol-chloroform extracted. The first enzyme cuts near the T3 promoter, downstream of the plasmid insert; this DNA was used with T7 RNA polymerase to produce a mouse *RPGRIP* sense transcript terminating at the *Hind*III cut. *Xho*I cuts on the other side of the insert so this DNA was used with T3 RNA polymerase to produce the antisense transcript. The RNA probes were made using the DIG RNA Labeling Kit (SP6/T7) (Boehringer Mannheim). For the sense transcript the T7 RNA polymerase supplied with the kit was used; for the antisense transcript a separate T3 RNA polymerase and its buffer (containing 0.4 M Tris-HCl, pH 8.0, 60 mM $MgCl_2$, 100 mM dithiothreitol, 20 mM spermidine; Boehringer Mannheim) were used. The RNA probes were ethanol precipitated in the following way: 2.5 µl of 4 M lithium chloride and 75 µl ethanol (at -20°C) were added to the standard labeling reactions (produced according to the DIG RNA Labeling Kit manufacturer's instructions), then mixed and incubated at -70°C for 30 minutes or -20°C for 2 hours. They were then centrifuged at 12,000 x g at 4°C for 15 minutes and the pellets washed carefully with 50 µl of 70% ethanol (at -20°C). The tubes were centrifuged in the same way again and the pellets dried under vacuum before being dissolved in 100 µl sterile RNase-free H_2O.

12.5 and 15 day old mouse embryo sections were supplied by Judy Fletcher (MRC Human Genetics Unit, Edinburgh). These were defrosted at room temperature for 1 hour. The riboprobes were diluted 1:200 in hybridisation buffer and vortexed before being denatured for 5-10 minutes at 70°C. 100 µl of each probe mix was added to the appropriate slide which was then overlayed with a coverslip. The slides were incubated overnight at 65°C in a sealed perspex box with 2 sheets of Whatman paper wetted with 1 x salt/50% formamide. After hybridisation, the slides were transferred to a coplin jar and washed for 15 minutes at 65°C in solution A to allow the coverslips to fall off. They were washed again, twice, in solution A for 30 minutes at 65°C, and twice more for 30 minutes at room temperature in 1 x TBST. The slides were blocked in 10% heat-inactivated sheep serum in 1 x TBST for at least 1 hour at room temperature. They were dried in the regions of the sections and then 100 µl 1:2000 anti-DIG AP-Fab fragment in 10% heat-inactivated sheep serum in 1 x TBST was added. The slides were incubated overnight in a humidified chamber at 4°C. They were washed 4 to 5 times for 20 minutes in 1 x TBST at room temperature and then twice for 10 minutes in 1 x NTMT at room temperature. The slides were stained in a coplin jar in the dark (i.e. wrapped in foil) at room temperature with 4.5 µl NBT plus 3.5 µl BCIP/1 ml NTMT. The staining reaction was checked after a few hours

and then left to proceed overnight at 4°C. The reaction was stopped by washing the slides twice in distilled water. The sections were fixed in 4% paraformaldehyde/0.1% glutaraldehyde for 20 minutes and then dehydrated through an alcohol series (30 seconds each wash). They were counterstained with filtered 0.1% eosin in 95% ethanol for 20 to 30 seconds and then rinsed twice for 30 seconds in 95% ethanol and washed twice for 5 minutes in 100% ethanol. They were transferred to xylene for 2 x 5 minutes and then mounted in DPX mounting medium (Fisher Scientific, Manchester, UK).

2.24 - Two hybrid library transformation into bacteria and amplification

The two hybrid retinal library constructed by C.H. Sung in the pACTII vector (Kumar *et al.*, 1996) contained 2 x 10^6 independent clones and was provided as 10 µl of DNA at a concentration of 66 ng/µl. A 40 µl aliquot of electro-competent Electromax DH10B *E. coli* cells was transformed with 300 ng of library DNA by electroporation. 1 ml of pre-warmed LB was added to the cells which were incubated in a rotary shaker at 37°C for 1 hour. Twenty 50 µl aliquots of this recovery culture were added to twenty 500 µl aliquots of LB and spread onto twenty 22 cm square L-agar plates containing 50 µg/ml ampicillin. At the same time, 0.01 µl of one 500 µl aliquot was spread onto a small L-agar plates containing 50 µg/ml ampicillin to assess transformation efficiency. The plates were incubated at 30°C for 24 hours instead of 37°C (on the advice of the commercial manufacturer (Clontech) of the pACTII vector) because at the higher temperature there was a risk that a temperature sensitive λ repressor would be inactivated leading to lysis by residual λ phages. Using the transformation control plate the total number of transformants was estimated to be 2.91 x 10^7. Several pastettes of LB medium containing 50 µg/ml ampicillin were added to each plate and the bacterial colonies scraped off and mixed into 2 litres of LB medium containing 50 µg/ml ampicillin. This was incubated at 30°C for 2 hours. After this, 5 ml was removed and used to make glycerol stocks and the rest used to prepare plasmid DNA.

2.26 - Yeast Two-Hybrid System

2.26.1 - Strain, media and solutions

Y190 genotype: *MATa, ura3-52, his3-200, ade2-101, lys2-801, trp1-901, leu2-3, -112, gal4Δ, gal80Δ, cyh^r 2, , LYS2 :: GAL1*$_{UAS}$ *-HIS3*$_{TATA}$*-HIS3, URA3::GAL1*$_{UAS}$ *- GAL1* $_{TATA}$ *– lacZ.* Yeast two-hybrid experiments were carried out in this *S. cerevisiae* strain, which

contains two GAL4-responsive reporter genes, *HIS3* and *lacZ*. The genes encoding GAL4 and its negative regulator gal80 have been deleted.

10 x Dropout solution

L-isoluecine	150 mg
L-Valine	750 mg
L-Adenine hemisulphate salt	100 mg
L-Arginine HCl	200 mg
L-Histidine HCl monohydrate	100 mg
L-Leucine	500 mg
L-Lysine	150 mg
L-Methionine	100 mg
L-Phenylalanine	250 mg
L-Threonine	1000 mg
L-Tryptophan	100 mg
L-Tyrosine	150 mg
L-Uracil	100 mg

dH_2O to 500 ml. Autoclave and store at $^{-}4^{\circ}C$

The following 10 x dropout solutions were prepared by omitting the appropriate amino acids in each case: 10 x dropout minus leucine and tryptophan, 10 x dropout minus histidine, leucine and tryptophan, 10 x dropout minus leucine, 10 x dropout minus tryptophan, 10 x dropout minus leucine and histidine, 10 x dropout minus tryptophan and histidine. These solutions were ingredients of the synthetic complete media that were used to select for yeast transformants containing plasmids carrying prototrophic genes.

Synthetic complete Medium

3.35 g Yeast Nitrogen Base without amino acids (Sigma) was mixed with 10 g agar (Difco). dH_2O was added to 425 ml, followed by 50 ml of the appropriate 10x dropout solution and the pH adjusted to 5.8. This was autoclaved and allowed to cool to approximately 55°C before the addition of 25 ml of a sterile 40% solution of glucose. If required, 3-amino-1,2,4-triazole was then added from a 1 M stock solution to a final concentration of 25 mM.

YPD/litre

Peptone (Difco)	20 g
Bacto yeast extract (Difco)	10 g
dH_2O	to 1 litre

Adjust to pH 5.8, autoclave, and cool to approximately 55°C. Add glucose to 2% (50 ml of a sterile 40% stock solution). For agar plates, add 18 g/litre agar (Difco).

10 xTE buffer

0.1 M Tris-HCl, 10 mM EDTA, pH 7.5. Autoclave.

10 x Lithium Acetate (LiAc)

1 M lithium acetate. Adjust to pH 7.5 with dilute acetic acid and autoclave.

50% PEG

50% PEG 3350. Autoclave.

PEG/LiAc solution

1 ml 10x TE

1 ml 10x LiAc

8 ml 50% PEG

Z Buffer:

$Na_2HPO_4.7H_2O$	8.05 g
$NaH_2PO_4.H_2O$	2.75 g
KCl	0.375 g
$MgSO_4.7H_2O$	0.123 g

dH_2O up to 500 ml. Adjust pH to 7.0

Z buffer/X-gal solution

100 ml Z buffer

0.27 ml β-mercaptoethanol

1.67 ml X-gal stock solution (20 mg/ml)

Salmon sperm DNA

The salmon sperm DNA (Sigma type III sodium salt) was dissolved in water to a concentration of 10 mg/ml and the concentration of NaCl adjusted to 0.1 M. The solution was extracted once with phenol and once with phenol-chloroform and the DNA sheared by passing it through a hypodermic needle 12 times. The DNA was precipitated by adding 2 volumes of ice cold ethanol, recovered by centrifugation and then redissolved in water at 10 mg/ml (the OD_{260} of the solution was measured and the exact concentration calculated). The DNA was placed in a boiling water bath for 10 minutes and then stored in small aliqouts at -20°C. Before using the salmon sperm DNA it was boiled for 5 minutes and then quickly chilled in ice water.

2.26.2 - Yeast Transformation Protocol

The amounts shown are for small scale transformations, those shown in parentheses are for library scale transformation.

1. 1 ml of YPD was inoculated with several large yeast colonies which were no more than 3 weeks old. The tube was vortexed to disperse clumps.
2. This was transferred to a flask containing 50 ml (150 ml) of YPD.
3. The flask was incubated overnight at 30°C with shaking at 250 rpm until the culture reached stationary phase (OD_{600} >1.5).
4. Sufficient overnight culture was transferred into 300 ml (1L) of YPD to produce an OD_{600} of 0.2-0.3.
5. This was incubated at 30 °C with shaking for 3 hours then the cells were centrifuged at 1,000g for 5 minutes at room tempreature.
6. The supernatant was removed and the cells resuspended in 50 ml (500 ml) dH_2O before centrifuging again.
7. The supernatant was discarded and the pellet resuspended in 1.5 ml (8 ml) of freshly prepared 1x TE/LiAc.
8. For each separate tranformation a separate tube was prepared containing 0.1 µg (0.2-1.0 mg) DNA-binding domain construct, 0.1 µg (0.1-0.5 mg) activation domain construct and 0.1 mg (20 mg) sheared salmon sperm DNA.
9. 0.1 ml (8 ml) of the competent yeast cells prepared in step 7 were added to each tube.
10. 0.6 ml (60 ml) sterile PEG/LiAc solution was added and the tubes vortexed to mix.
11. The tubes were incubated at 30 °C with shaking for 30 minutes.

12. 70 µl (7.0 ml) DMSO was added to each tube and mixed by gentle inversion.
13. The tubes were placed in a 42 °C water bath for 15 minutes to heat shock. The cells were gently mixed occasionally in the case of library scale transformations.
14. The cells were chilled on ice and then centrifuged for 5 seconds (5 minutes) at 12,000 rpm (1000 x g).
15. The pelleted cells were resuspended in 100 µl (10 ml) 1x TE.
16. For small scale transformations, half the cells were plated onto synthetic complete medium lacking tryptophan and leucine, and half onto synthetic complete medium containing 25 mM 3-aminotriazole and lacking tryptophan, leucine and histidine. For library scale transformations, 100µl of 1:1000, 1:100 and 1:10 dilutions were spread onto synthetic complete medium lacking tryptophan and leucine. The colonies that grew on these plates were counted and used to estimate transformation efficiency. The remainder was spread over sixteen 22cm square plates of synthetic complete medium containing 25 mM 3-aminotriazole and lacking tryptophan, leucine and histidine.
17. The plates were incubated at 30 °C until colonies became visible.

2.26.3 - β-Galactosidase colony-lift filter assay

Fresh colonies were used for the β-Galactosidase assay (i.e. no more than 1 week old and between 1 and 3 mm in diameter). Z buffer/X-gal solution was prepared as described above and used to soak clean pieces of Schleicher and Schuell (London, UK) GB003 gel blotting paper which had been cut to fit inside 8 cm petri dishes. Pieces of Hybond (Amersham Pharmacia Biotech) N+ nucleic acid transfer membrane of the same size were cut and placed onto the surfaces of the plates of colonies to be assayed. The membranes were pierced in three asymmetric positions while still in contact with the colonies in order to orient them with the agar plates. Using forceps, the membranes were removed from the agar and tranferred to a liquid nitrogen bath where they was submerged, colony side up, for 10 seconds. They were removed from the liquid nitrogen bath and allowed to thaw at room temperature (the freeze/thaw treatment permeabilises the yeast cells). The membranes were placed onto pieces of pre-soaked gel blotting paper and transferred to a 30°C incubator. They were periodically checked for the appearance of blue colonies until 8 hours had elapsed. Blue colonies appearing after 8 hours were deemed to be false-positives.

2.26.4 - Plasmid purification from transformed yeast colonies

This method is based on that of Hoffman and Winston (1987). 200 µl of plasmid release buffer (2% Triton X 100 (Sigma), 1% SDS, 0.1 M NaCl, 0.01 M Tris pH 8.0, 0.001 M EDTA) was put into a 1.5 ml microfuge tube a 50 µl volume of yeast from a streaked plate was added. 300 µl glass beads (Sigma), 100 µl phenol and 100 µl chloroform were added. The tube was vortexed for 2 minutes and then centrifuged at 12,000 rpm at room temperature for 5 minutes. The supernatant was removed to a new tube and 200 µl of chloroform was added. The tube was vortexed and centrifuged as before. The supernatant was removed and 25 µl 5 M NaCl and 1 ml absolute ethanol were added. The tube was placed at -70°C for 15 minutes and then centrifuged at 12,000 rpm at room temperature for 15 minutes. The pellet of DNA was washed in 70% ethanol, air dried and resuspended in 25 µl dH_20.

2.26.5 - Transformation of *E. coli* with plasmid isolated from yeast.

Plasmid DNA preparations from yeast are of low concentration and contain yeast genomic DNA. For these reasons, *E. coli* were transformed with these preparations by electroporation because of the relatively high transformation efficiency that can be achieved. The HB101 strain carrying the *leuB* mutation was used so that transformants receiving the pACTII plasmid (and not the pAS1 plasmid) could be selected by complementation with the yeast *LEU2* gene.

Electrocompetent HB101 cells were prepared as follows:

1. 5 ml of LB was inoculated with a single HB101 colony and incubated at 37°C, with shaking, overnight.
2. 1 ml of the overnight culture was transferred to a 500 ml flask containing 100 ml LB medium. This was incubated at 37 °C until the OD_{600} reached 0.5 (±0.03).
3. The flask was chilled on ice for 20 minutes and the culture was then centrifuged at 1200 x g for 5 minutes at 4°C.
4. The cells were resuspended in 100 ml of an ice-cold solution of 10% glycerol and left on ice for 20 minutes.
5. The cells were centrifuged at 1200 x g for 10 minutes at 4°C. The supernatant was removed and the cells gently resuspended in 10 ml ice-cold 10% glycerol and incubated on ice for 20 minutes.

6. The cells were again centrifuged at 1200 x g for 10 minutes at 4°C and resuspended in 400 µl of ice-cold 10% glycerol. At this point the cells were either electroporated or snap-frozen and stored at -70°C.

Samples of plasmid DNA isolated from yeast were used to transform 40 µl aliquots of electrocompetent HB101 cells. After recovery, the cells were washed twice in M9 minimal medium and plated onto M9 agar medium containing 50 µg/ml ampicillin and a 1 in 10 dilution of 10 x dropout solution minus leucine. The plates were incubated at 37°C overnight. Plasmid DNA was isolated from Leu^+, Amp^r HB101 transformants using the standard mini-prep procedure.

2.27 - Preparation of *in vitro* translated proteins

To prepare *in vitro* translated proteins the TnT quick coupled transcription/translation system was used (Promega). This system provides the "TnT quick master mix", which combines the reagents for transcription and translation and so allows these processes to be carried out in a single tube. The mix contains RNA polymerase, nucleotides, salts and a ribonuclease inhibitor for transcription, and rabbit reticulocyte lysate for translation. To synthesize proteins, the mix is incubated with a template DNA and a source of radioactive methionine. Reactions are set up in a 0.5 ml thin-walled microcentrifuge tube by combining 20 µl of the TnT quick master mix with 1 µl 1 mM [^{35}S] methionine (1,000 Ci/mmol at 10 mCi/ml; Amersham Pharmacia Biotech), 1 µg DNA template and nuclease-free water to a final volume of 25 µl. The components are mixed gently by pipeting and the tube is incubated at 30°C for 90 minutes. An aliquot of the reaction is analysed on an SDS-polyacrylamide gel.

2.28 - SDS-polyacrylamide gel electrophoresis

SDS-polyacrylamide gels were used to separate mixtures of proteins and compare migration of individual polypeptides with the migration of molecular weight standards. A Bio-Rad Protean II mini gel apparatus was used. The glass plates were assembled in the gel former and the resolving gel poured. 5 ml of 12% resolving gel contained 1.6 ml of dH_2O, 2.0 ml of 30% acrylamide mix (Severn Biotech, Kidderminster, UK), 1.3 ml of 1.5 M Tris (pH 8.8), 0.05 ml of 10% SDS, 0.05 ml of freshly prepared 10% ammonium persulfate and 2 µl

TEMED (Gibco BRL). Approximately 1 ml of TE-saturated butanol was layered on top of the resolving gel solution. This was immiscible with the gel solution and so remained on top causing the gel to set straight and also prevented oxygen from diffusing into the gel and inhibiting polymerisation. After the resolving gel had set (45 minutes) the butanol was discarded and the top of the gel washed in water and dried with a paper towel. A comb was placed above the resolving gel and the stacking gel poured around the comb. 2 ml of stacking gel contained 1.4 ml of dH_2O, 0.33 ml of 30% acrylamide mix, 0.25 ml of 1.0M Tris-HCl (pH 6.8), 0.02 ml of 10% SDS, 0.02 ml of freshly prepared 10% ammonium persulfate and 2 µl TEMED. The stacking gel was allowed to set for approximately 30 minutes before the comb was removed and the wells rinsed with dH_2O to remove any unpolymerised acrylamide. The gel was mounted in the electrophoresis apparatus and tri-glycine buffer added to the top and bottom reservoirs. Tris-glycine buffer is made by diluting a 5x stock solution in dH_2O. The 5x stock is prepared by dissolving 15.1 g of Tris base and 94 g of glycine in 900 ml dH_2O, adding 50 ml of 10% (w/v) electrophoresis-grade SDS and adjusting the volume to 1000 ml with dH_2O. Samples are prepared by mixing them with an equal volume of 2x SDS-loading buffer and heating for 3 minutes in a boiling water bath. 2x SDS-loading buffer contains 100 mM Tris-HCl (pH 6.8), 100 mM dithiothreitol, 4% SDS, 0.2% bromophenol blue and 20% glycerol. Samples are loaded into the bottom of the wells. The electrophoresis apparatus is connected to a power supply and a voltage of 200 volts applied to the gel for 55 minutes (mini-gels).

2.29 - Co-immunoprecipitation

The Matchmaker Co-IP kit (Clontech) was used to co-immunoprecipitate *in vitro* translated epitope-tagged RPGR and RPGRIP proteins. The T7 promoter and c-myc epitope tag were introduced upstream of the *RPGR* cDNA by amplifying the pAS1.RPGR insert from the yeast two-hybrid bait construct by PCR using the DNA-binding domain primers supplied with the kit (the first underlined portion is the T7 promoter, the translation initiation codon is in bold and the second underlined portion encodes the c-myc epitope tag):

DNA-Binding domain (BD) forward Co-IP primer
5'-AAAATTGTAATACGACTCACTATAGGGCGAGCCGCCACC**ATG**GAGGAGCAGA
AGCTGATCTCAGAGGAGGAGCTGGGTCAAAGACAGTTGACTGTATCG-3'

DNA-BD reverse Co-IP primer

5'-TACCTGAGAAAGCAACCTGACCTACAGG-3'

The T7 promoter and haemagglutinin (HA) epitope tag were introduced ustream of the *hRPGRIP* cDNA by amplifying the pACTII.RPGRIP insert from the yeast two-hybrid prey construct using the activation domain primers supplied with the kit (the first underlined portion is the T7 promoter, in bold is the translation initiation codon, the second underlined portion encodes the HA epitope tag):

activation domain (AD) forward Co-IP primer

5'-

AAAATTGTAATACGACTCACTATAGGGCGAGCCGCCACC**ATG**TACCCATACGA CGTTCCAGATTACGCTCCACCAAACCCAAAAAAAGAG-3'

AD reverse Co-IP primer

5'-ACTTGCGGGGTTTTTCAGTATCTACGAT-3'

The forward primers anneal to the plasmids at the multiple cloning site, i.e. between the GAL4 domain-coding sequence and the *RPGR/RPGRIP* coding sequence. This means that only the bait/library proteins are translated and not the GAL4 domains.

PCR mixes were made up with the following components:

36.25 µl	dH_2O
5 µl	10x Expand HiFidelity PCR buffer (Boehringer Mannheim)
1 µl	DNA template (100 ng/µl)
1 µl	DNA-BD/AD forward Co-IP primer (10 µM)
1 µl	DNA-BD/AD reverse Co-IP primer (10 µM)
5 µl	dNTP mix (2 mM each)
0.75 µl	Expand HiFidelity Polymerase mix (Boehringer Mannheim).

The reactions were run in a MJ Research PTC-200 Peltier thermal cycler using the following program: 94°C for 1 minute; 2 cycles of 94°C for 15 seconds, 72°C for 5 minutes; 2 cycles of 94°C for 15 seconds, 70°C for 5 minutes; 21 cycles of 94°C for 15 seconds, 68°C for 5 minutes; then 68°C for 7 minutes.

5 µl of the reaction was analysed by electrophoresis on a 1.2% agarose gel containing 1 mg/ml ethidium bromide. The remaining 45 µl was purified by phenol-chloroform extraction.

The purified PCR products were used with the TnT Quick Coupled Transcription/Translation System (Promega) to prepare [^{35}S] methionine-labelled proteins (see section 2.27).

To perform the co-immunoprecipitation, 5 µl *in vitro* translated bait protein (RPGR) was mixed with 5 µl *in vitro* translated library protein (RPGRIP) in a 1.5 ml microcentrifuge tube. These were mixed gently with a pipette and incubated at 30°C for 1 hour with occasional mixing. 470 µl co-immunoprecipitation buffer was added (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM DTT, 5 µg/ml aprotinin, 0.5 mM PMSF, 0.1% Tween 20) and then 10 µl protein-G agarose beads (a 50% slurry in Tris buffered saline containing 0.1% Tween 20; Clontech) and 10 µl c-myc monoclonal antibody (mouse IgG, 0.1 µg/µl; Clontech) or 10 µl HA polyclonal antibody (affinity purified rabbit Ig, 0.1 µg/µl; Clontech). The tube was incubated at 4°C for 2 hours with continuous rocking. The tube was centrifuged at 14,000 rpm for 1 minute and the supernatant discarded. 0.5 ml Tris buffered saline containing 0.1% Tween 20 was added (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Tween 20) and the tube agitated gently. The tube was centrifuged again and the supernatant discarded. The protein-G agarose beads were washed three times in this way. After the third wash the supernatant was discarded and the beads resuspended in 15 µl SDS-loading buffer (100 mM Tris-HCl pH 6.8, 100 mM dithiothreitol, 4% SDS, 0.2% bromophenol blue and 20% glycerol). The samples were eluted and denatured by heating to 80°C for 5 minutes. The tube was placed on ice and then centrifuged brielfy. 10 µl of the supernatant was loaded onto a sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) minigel. Following electrophoresis the gel was transfered to a tray containing gel fixation solution (20% methanol, 10% acetic acid) and placed on a rotary shaker for 10 minutes at room temperature. The solution was replaced with fresh solution and the gel soaked for 30 minutes. The gel was rinsed in dH_2O and then soaked in Amplify Fluorographic Reagent (Amersham Pharmacia Biotech) on a rotary shaker for 20 minutes at room temperature. The gel was then placed onto pre-wetted Whatman 3MM paper, covered with Saran wrap and dried for 1 hour at 80°C under constant vacuum in a Bio-Rad model 583 gel dryer. The dried gel was exposed to Biomax X-ray film (Kodak) overnight and the film developed in a Fuji RGII X-ray film processor.

2.30 - Mutation screening

Frozen lymphoblastoid cell lines from the three affected individuals (BB13, BB15 and BB39) were thawed and re-established by Alan Lennon (MRC Human Genetics Unit). Total RNA was purified from cell pellets and used to amplify RPGRIP cDNA (in two overlapping fragments) by RT-PCR. Two rounds of PCR using nested or hemi-nested primers were required to amplify the RPGRIP cDNA using the following oligonucleotides (the sequences of which are given in table 2.1, page 59):

BB13 5' fragment	26Fwd and 19 Rev (first round)
	-6Fwd and RG32 Long Rev (second round)
BB15 5' fragment	-6Fwd and 19Rev (first round)
	-6Fwd and RG32 Long Rev (second round)
BB39 5' fragment	-6Fwd and 44Rev (first round)
	-6Fwd and 19Rev (second round)
BB13/15/39 3' fragments	1153 Fwd and Stop+36 (first round)
	1153Fwd and Stop+6 (second round).

Sequence changes detected in the patient RPGRIP cDNAs were inspected at the genomic level by PCR-amplifying and sequencing fragments of genomic DNA in the region of the changes using the primers detailed in Table 2.2.

Table 2.2: PCR primers used to amplify and sequence RPGRIP genomic DNA in the regions of variations from wild-type detected in RPGRIP RT-PCR products (primer sequences given in Table 2.1).

DNA	Primers used to amplify RPGRIP genomic DNA	Primer used to sequence genomic PCR product
BB13	869 Fwd and Intron 5 Rev	869 Fwd
	Intron 5 Fwd and 962Rev	Intron 5 Fwd
	Intron 3 Fwd and 670 Rev	Intron 3 Fwd
BB15	869 Fwd and Intron 5 Rev	869 Fwd
	Intron 5 Fwd and 962Rev	Intron 5 Fwd
	Intron 3 Fwd and 670 Rev	Intron 3 Fwd
BB39	686 Fwd and Intron 4 Rev; Intron 4 Fwd and 845 Rev	686Fwd, Intron 4 Rev and Intron-4 Fwd
	1048 Fwd and Intron 6 Rev	Intron 6 Rev

CHAPTER THREE

THE YEAST TWO-HYBRID SYSTEM

3.1 - INTRODUCTION

The yeast two-hybrid system (Fields, and Song, 1989; Chien *et al.*, 1991) allows the identification of interacting partners for a gene of interest by exploiting the modular nature of transcription factors. Two potential interacting proteins are expressed in *S. cerevisiae*, one of which contains the protein of interest ("bait", X) fused to the DNA-binding domain (BD) of a yeast transcription factor such as GAL4, while the other ("prey", Y) is fused to the corresponding activation domain (AD). Binding of the bait to an interacting partner leads to the restoration of a functional transcription factor, which activates the GAL4 upstream activation sequence (UAS) and downstream reporter genes such as *HIS3* and *lacZ* (Figure 3.1).

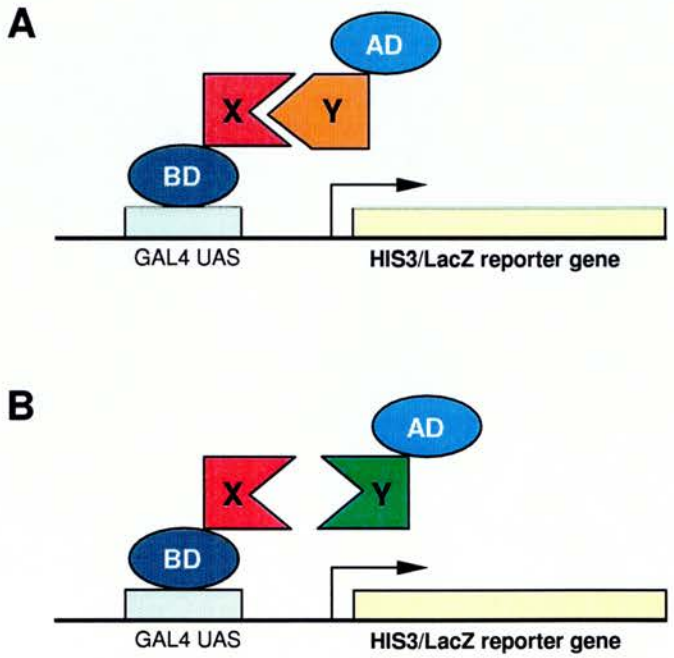


Figure 3.1: Schematic representation of the yeast two hybrid system. A: A fusion protein consisting of the GAL4 DNA-binding domain (BD) and protein 'X' occupies the GAL4 UAS. Protein 'Y' interacts with 'X' and brings the GAL4 activation domain (AD) into contact with BD, so restoring a functional transcription factor. This leads to transcription of the *HIS3* and *LacZ* reporter genes. B: Protein 'X' does not interact with protein 'Y', the GAL4 transcription factor components are kept separate and the reporter genes are not transcribed.

A variation of this is the yeast one-hybrid system, which allows expression libraries to be screened for proteins that directly interact with DNA target sequences. Yeast cells are transfected with 2 plasmids, the first of which consists of the transcription factor activation domain cDNA fused to the cDNA library insert. The second contains the DNA binding site under investigation plus the GAL1 promoter in front of the *HIS3* gene. Binding of the protein encoded by the library plasmid to the DNA target sequence under investigation then leads to the initiation of *HIS3* reporter gene expression (Li and Herskowitz, 1993). The reverse two-hybrid and one-hybrid systems are also modifications of the two-hybrid system, whereby maintenance of an interaction between two proteins, or between a DNA binding protein and its target sequence, is deleterious to growth. These are used to identify mutations or additional molecules that disrupt interactions between two proteins or a DNA-binding protein and its target sequence (Vidal *et al.*, 1996).

The GAL4 yeast transcriptional activator is composed of two physically separable, functionally independent domains: a DNA-binding domain and a transcriptional activation domain (Keegan, Gill and Ptashne, 1986; Hope and Struhl, 1986; Ma, and Ptashne, 1987). The DNA-binding domain recognises and binds to upstream activating sequences (UAS), and the transcriptional activation domain interacts with other components of the transcription machinery required to initiate transcription. In yeast, UASs comprise one type of *cis*-acting transcription element. They are recognised by specific transcriptional activators and enhance transcription from adjacent downstream TATA boxes. An example of this is found in the control of galactose metabolism, which is controlled by the GAL4 and GAL80 proteins, and the carbon source in the growth medium (Guthrie and Fink, 1991). When galactose is present, the GAL4 protein binds to GAL4-responsive elements in the UAS upstream of several genes involved in galactose metabolism, including *GAL1*, and activates transcription. In the absence of galactose, the GAL80 protein binds to GAL4 and blocks transcriptional activation. In the presence of glucose, transcription of the galactose genes is immediately repressed (Johnston, Flick and Pexton, 1994). All of the UASs of the 20 known galactose-responsive genes contain one or more conserved palindromic sequences to which the GAL4 protein binds (Giniger, Varnum, and Ptashne, 1985). These are 17-mers and function in an additive way, that is, the occupancy of more sites leads to higher levels of transcription (Giniger, and Ptashne, 1988).

In the yeast two-hybrid system used in this study, two potentially-interacting proteins are expressed in *S. cerevisiae* strain Y190 from the pAS1 and pACTII cloning vectors. One of the proteins is fused to the GAL4 DNA-binding domain (bait; pAS1), and

the other is fused to the GAL4 transcriptional activation domain (prey; pACTII). The recombinant hybrid proteins are targeted to the nucleus; for the pAS1 fusion protein the nuclear localisation signal is an intrinsic part of the GAL4 DNA-binding domain (Silver, Keegan, and Ptashne, 1984) whereas the pACTII fusion protein uses the SV40 large T antigen nuclear localisation signal. If the proteins interact, the DNA-binding domain will be tethered to the transcriptional activation domain. The GAL4 transcriptional activator will thus be reconstituted and will activate transcription of the *lacZ* and *HIS3* reporter genes. In GAL4 yeast two-hybrid systems, the native *GAL1* UAS or a synthetic UAS_G 17-mer consensus sequence provides the binding sites for the GAL4 DNA binding domain. The *lacZ* and *HIS3* reporter genes of Y190 have been engineered to utilise the *GAL1* UAS. The Y190 yeast strain also has deletions of endogenous *gal4* and *gal80* genes to prevent interference with two-hybrid system constructs and to prevent glucose repression of reporter gene activity.

The Y190 *HIS3* reporter gene is under the control of the *GAL1* UAS and a minimal promoter containing two TATA boxes from the promoter of the native yeast *HIS3* gene. One of the TATA boxes, referred to as Tc, drives low-level constitutive expression of the *HIS3* gene (Mahadevan, and Struhl, 1990). A positive two-hybrid interaction will induce high-level expression because of binding to the *GAL1* UAS. However, in the absence of an interaction a significant level of *HIS3* expression is promoted by the Tc TATA box. To overcomes this residual *HIS3* expression, cells must be grown in the presence of 3-aminotriazole (3-AT) a chemical inhibitor of the *HIS3* gene product (Durfee *et al.*, 1993).

Expression of the fusion proteins in yeast is driven by the *ADH1* promoter. Expression is high for most of the logarithmic growth phase of the host cells, but is repressed late in the logarithmic phase by the ethanol that accumulates in the growth medium (Ammerer, 1983). The pACTII and pGAD10 vectors contain a truncated *ADH1* promoter, which is capable only of driving much lower levels of expression (Ruohonen, Aalto, and Keranen, 1995). In pACTII however, a portion of DNA from pBR322 coincidentally acts as a transcriptional enhancer and produces strong constitutive expression (Tornow and Santangelo, 1990).

3.2 – ALTERNATIVE APPLICATIONS OF THE YEAST TWO-HYBRID SYSTEM

The yeast two-hybrid system can firstly be used to test specific candidate proteins for an ability to interact with the protein of interest. The candidates may be chosen on the basis of physiological experiments, biochemical or genetic data, sequence similarity, or other criteria.

A second approach is to use the yeast two-hybrid system to screen entire cDNA libraries for interacting proteins. Here, the principal choice concerns the tissue source of the library being screened (for example, to search for interacting partners for a human retinal protein the best library would be produced from human retina RNA). Two-hybrid libraries are constructed in the activation domain plasmid (rather than the DNA binding domain plasmid) because of the abundance of sequences capable of activating transcription when fused to a DNA binding domain (Ma, and Ptashne, 1987). In this project, the yeast two-hybrid system was used both to assess the ability of candidate proteins to interact with RPGR and to screen cDNA libraries for RPGR-interactors.

3.3 - CANDIDATE RPGR-INTERACTING PROTEINS

At the start of this project, little was known about the function of the RPGR protein. The only clues as to the cellular role of RPGR came from expression studies, sequence homology searches and the nature of the XLRP disease process itself. As described in Chapter 1, the retina contains a large number of diverse proteins capable of causing a retinitis pigmentosa phenotype when mutated. On the basis of the clinical phenotype, the RPGR protein could be involved in many different cellular processes in the retina. Alternatively, it could be a component of an as yet unidentified pathway. Expression studies have indicated that the *RPGR* gene is expressed in all of the tissues examined yet the disease only affects the retina (Meindl *et al.*, 1996). It is possible that the role of RPGR in the retina is distinct from other tissues, and that it interacts with a protein that is retina-specific. In non-retinal tissues, there may be a related protein that compensates for the mutant RPGR in XLRP patients. This is speculation however, and the expression pattern tells us little about the function of RPGR. The most salient feature of the DNA sequence is the RCC1-homologous domain spanning exons 2 to 10. The candidate proteins that were tested for an ability to interact with RPGR were chosen on the basis of this homology, and included Ran, ARF1 and Clathrin heavy chain.

3.3.1 - Ran

Ran (Ras-related nuclear protein) is a small GTPase for which RCC1 acts as a guanine nucleotide exchange factor (Bischoff and Ponstingl, 1991). The structural similarity between the RPGR N-terminal domain and RCC1 (Renault *et al.*, 1998) led to the proposal that this RPGR domain also interacts with Ran.

Ran is a 25 kDa protein consisting of 216 amino acids and is a member of the Ras superfamily of small GTP-binding proteins. It was originally isolated by screening a human cDNA library for sequences related to the Ras family (Drivas *et al.*, 1990). Homologues of human Ran have been identified in mammals, birds, *Drosophila, Caenorhabditis elegans,* plants, *Plasmodium falciparum, Dicytostelium discoideum, Giardia lamblia, Schizosaccharomyces pombe* and *Saccharomyces cervisiae* (Bischoff *et al.*, 1996). They are all located predominantly in the nucleus and share a unique, highly acidic (DEDDDL) carboxy-terminus (Bischoff *et al.*, 1996).

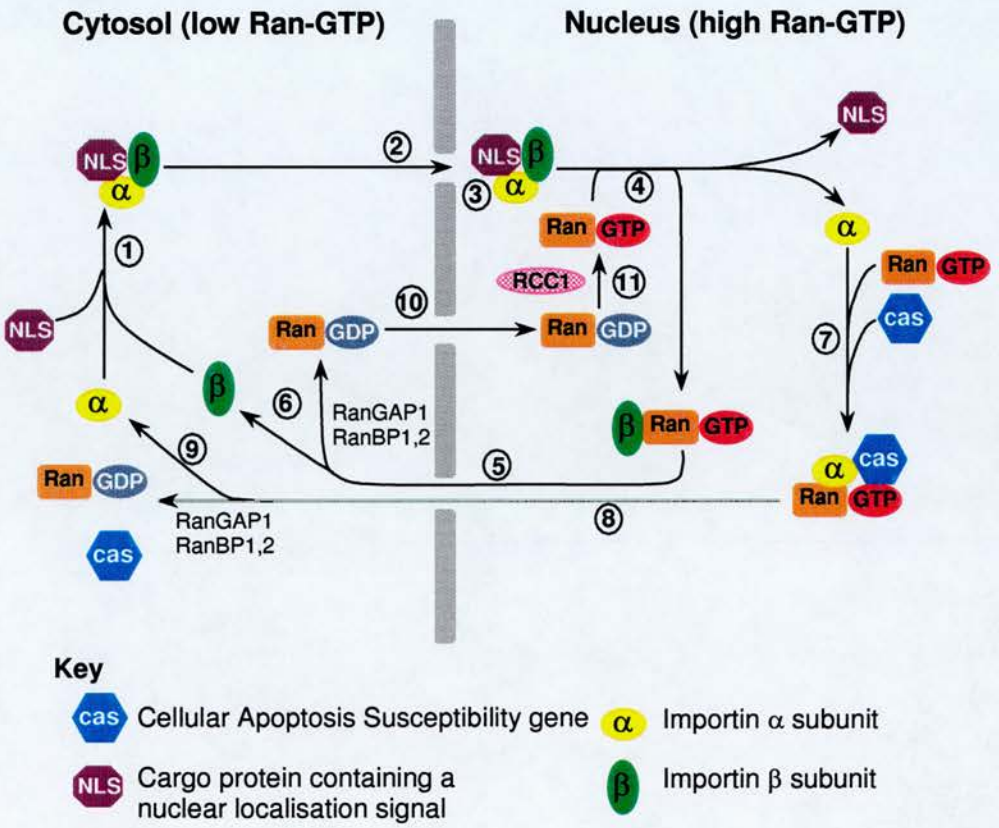


Figure 3.2: A model showing the involvement of Ran and RCC1 in NLS-mediated protein import. Steps 1 to 9 as follows: 1, formation of an import complex comprising importin α, importin β and the transport substrate ; 2, docking of the transport complex on the nuclear pore complex; 3, Ran- and energy-dependent translocation of the complex across the nuclear pore; 4, Ran-GTP binding to importin β and dissociation of transport complex; 5, export of importin β/Ran-GTP to the cytosol; 6, Ran-GTP hydrolysis, stimulated by RanBP1 and RanBP2-bound RanGAP1, and consequent dissociation of Ran-GDP from importin β; 7, Binding of Ran-GTP and CAS to importin α; 8, export of Ran-GTP/CAS/ importin α complex; 9, Ran-GTP hydrolysis, stimulated by RanBP1 and RanBP2-bound RanGAP1, and consequent dissociation of the Ran-GTP/CAS/ importin α complex; 10, Ran-GDP import; and 11, RCC1-dependent regeneration of Ran-GTP from Ran-GDP (Köhler, *et al.* 1999).

Ran has been implicated in various nuclear processes, including RNA metabolism, DNA replication and chromosome condensation and decondensation. However, most of the experimental data concerns its involvement in the organisation of microtubule structures during M phase of the cell cycle (reviewed by Kahana and Cleveland, 1999; Desai and Hyman, 1999; Nishimoto, 1999) and in the nucleocytoplasmic trafficking of proteins and RNA (reviewed by Moore, 1998; Kahana and Cleveland, 1999; Yoneda *et al.*, 1999), the latter being the better understood. The distributions of Ran-GTP and Ran-GDP in the cell, and the compartmentalisation of the GEF and GTPase Activating Protein (GAP) that generate them, are thought to underlie the directional nature of nuclear transport. RCC1, which promotes the formation of Ran-GTP, is localised to the nucleus while RanGAP1 and Ran-Binding Protein 1, which enhance the GTP hydrolysis activity of Ran, are exclusively cytoplasmic. Ran is found throughout the cell. Thus, there is an excess of Ran-GTP in the nucleus and Ran-GDP in the cytoplasm. Typically, Ran-GTP interacts with nuclear export receptors and guides transport of their cargoes to the cytoplasm and binds to import factors and causes dissociation of imported transport complexes and Ran-GDP associates with nuclear import receptors and directs movement into the nucleus (see Kahana and Cleveland, 1999). Figure 3.2 shows a simplified model of Ran-regulated nuclear protein import.

3.3.2 - ARF1

The ADP-ribosylation factor (ARF) family is the most divergent member of the Ras superfamily and consists of the ARF proteins and the ARF-like proteins (Kahn, 1996). The ARF proteins are essential components of a number of different vesicular trafficking pathways in eukaryotic cells and activators of specific forms of phospholipase D. Proteins were previously designated as ARF if they had the ability to (i) stimulate the ADP-ribosyltransferase activity of the cholera toxin A subunit and (ii) rescue *Saccharomyces cerevisiae* mutants with the double deletion of *ARF1* and *ARF2* genes (Moss and Vaughan, 1998). The ARF-like proteins are structurally similar to the ARF proteins and were initially thought to lack these ARF activities, although it has subsequently been shown that, under suitable conditions, this is not always the case (Hong *et al.*, 1998).

The mammalian ARF proteins are divided into three classes on the basis of sequence comparison. Class I consists of ARF1-ARF3 and is the best understood category. The Class I ARFs, which seem to be functionally redundant, regulate the assembly of several types of vesicle coat found at discrete steps in intracellular membrane transport (Roth, 1999). ARF1

is necessary for the formation of COP1 vesicles *in vitro* (Rothman and Wieland, 1996) and certain clathrin-coated vesicles both on endosomes and at the *trans*-Golgi network (Stamnes and Rothman, 1993). ARF1, like other members of the Ras superfamily, cycles between active GTP-bound and inactive GDP-bound forms. The conformational change that is induced by GDP/GTP exchange is coupled to membrane-binding through the ARF1 *N*-myristoyl moiety. Active ARF1-GTP is, thus, bound to intracellular membranes, and inactive ARF1-GDP is cytosolic (Goldberg 1998) (shown schematically in Figure 3.3). Exchange of GDP for GTP (and activation of ARF) is promoted by ARF GEFs, the majority of which contain the 200 amino acid Sec7 domain (Morinaga, *et al.*, 1997). RPGR lacks this domain. The functions of ARFs 4 and 5, which make up class II, are not known. ARF6, the least conserved of the human ARFs, constitutes class III and has been implicated in the regulation of membrane trafficking and remodelling of the plasma membrane and underlying cytoskeleton (Roth, 1999).

ARF1 interacts with the protein p532 (Rosa *et al.*, 1996a; Rosa *et al.*, 1996b), which was identified as a candidate RPGR-interactor because, like RPGR, p532 contains domains that are homologous to RCC1: RCC1-like domain 1 (RLD-1) and RCC1-like domain 2 (RLD-2). RLD1 shares 30% identity with RCC1 and 51% similarity (at the protein level); RLD2 shares 29% identity with RCC1 and 49% similarity. RLD-2 binds to ARF1 and RLD-1 stimulates its guanine nucleotide exchange activity.

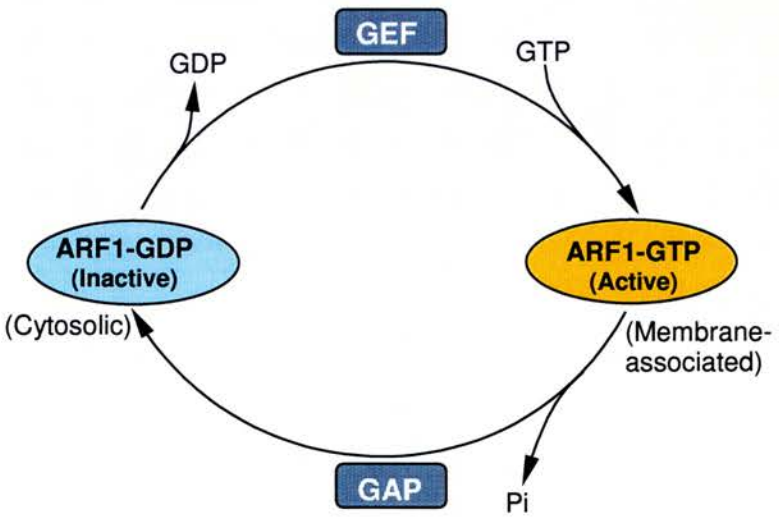


Figure 3.3: ARF1-GTP, generated by ARF1-specific GEFs, is the active form and membrane-bound. Inactive ARF1-GDP is cytosolic and is generated by ARF1-specific GAPs that promote GTP hydrolysis.

3.3.3 - Clathrin Heavy Chain

A third candidate interacting protein is the Clathrin heavy chain. This was selected as a candidate interactor because it was identified in a yeast two-hybrid screen for proteins that interact with the carboxy-terminal p532 RCC1-like domain (RLD-2) (Rosa and Barbacid, 1997). Clathrins are triskelion structures composed of three non-covalently bound heavy chains and 3 light chains, and are located on the cytoplasmic face of coated vesicles and coated pits. These intracellular organelles, of which clathrins constitute a major component, play a role in endocytosis and Golgi sorting (Robinson *et al.*, 1996).

Clathrin-mediated endocytosis is a means of efficiently retrieving proteins from the plasma membrane (reviewed by McMahon, 1999). This is important at presynaptic nerve terminals and in many cellular processes, including nutrient uptake and antigen presentation, and is a means of virus entry into cells. Three classes of proteins are known to contribute to the clathrin coat: transmembrane receptors, adaptor protein complexes and clathrin triskelia. Transmembrane proteins (for example, receptors for extracellular ligands) have, in most cases, internalisation motifs in the cytoplasmic domain that are recognised by the AP-2 adaptor complex proteins (or AP-1 at the *trans*-Golgi network), to which they become bound. This binding is followed by the recruitment of soluble clathrin from the cytoplasm, which also binds to adaptor complex proteins. The clathrin triskelia polymerise, under the influence of the adaptor protein complex, to form a cage-like structure that encloses the membrane vesicle. A lattice of pentagons and hexagons is formed by the polymerised clathrin and this induces curvature in the cage, so allowing membrane invagination. Another component of this pathway is the GTPase dynamin, which is required for coated pits to pinch off from the donor membrane as coated vesicles.

3.4 - TESTING CANDIDATES

Prior to the identification of a mutation hot-spot in *RPGR* exon ORF15, the majority of *RPGR* mutations associated with XLRP were clustered in the RCC1-like domain, suggesting that this is a functionally important domain. The yeast two-hybrid bait construct was therefore made to contain the first 12 exons of *RPGR*, which includes the whole of this RCC1-like domain.

Human *RPGR* exons 1-12, encoding amino acids 1 to 502, were amplified as a single fragment from an adult brain cDNA library. Ran and ARF1 were amplified from a human testis cDNA library (using the polymerase chain reaction). The primers used to

amplify these sequences (RPGRATGNcoI, Exon12BamHI, RanNcoI, RanStopBamHI, ARF1ATGNcoI, and ARFStopBamHI) had non-annealing 5' ends encoding restriction endonuclease cleavage sites (see Materials and Methods, page 58). These primer restriction sites facilitated the subcloning of the PCR products into the yeast two-hybrid vectors as *NcoI/Bam*HI fragments. The positioning of the restriction sites also ensured that the reading frame of each subcloned sequence was the same as the vector sequence encoding the GAL4 domain. All constructs were sequenced to ensure that errors had not been incorporated during the amplification reactions. The clathrin heavy chain construct, in the pGAD-GH GAL4 activation domain vector (van Aelst *et al.*, 1993), and the p532 RLD2 construct, in the pGBT9 GAL4 DNA binding domain vector (Bartel *et al.*, 1993a), were donated by Dr J.L.Rosa (University of Barcelona). Both of these two-hybrid vectors are compatible with the pAS2/pACTII system.

Competent Y190 cells were transfected with aliquots of the yeast two-hybrid system constructs and plated onto SC medium containing 3-AT, and lacking leucine, tryptophan and histidine. *HIS3*$^{+}$ colonies were tested for β-galactosidase activity.

Each of the three candidates was tested for an interaction with RPGR both as a fusion with the GAL4 DNA-binding domain and with the activation domain. This was done because certain interactions activate transcription more efficiently in one orientation (in terms of which protein is fused to which GAL4 moiety) than the other (Fields and Sternglanz, 1994).

3.4.1 - Positive Controls

The yeast proteins SNF1 and SNF4 are known to interact in the yeast two-hybrid system (Fields and Song, 1989). SNF1 is a serine-threonine-specific protein kinase, and SNF4 is a protein that is physically associated with SNF1 and required for its maximal activity. The genes encoding SNF1 and SNF4 in the pAS1 and pACTII vectors respectively (obtained from Mhairi Wallace, MRC Human Genetics Unit, Edinburgh) were transfected into the Y190 yeast strain as a control for a positive interaction. RPGR is known to interact with the cyclic GMP phosphodiesterase delta subunit (PDEδ) (Linari *et al.*, 1999). PDEδ was amplified by PCR from a human pancreas cDNA library using the primers PDEdNcoI and PDEdBamHI (see Materials and Methods, page 59). It was cloned into the pACTII vector and transfected with the RPGR bait plasmid into Y190. The interaction reported by Linari *et al.*, (1999) was verified here and provided a control for a positive interaction with RPGR. As another positive control, an aliquot of competent Y190 cells was transfected with yeast

two-hybrid bait and prey vectors containing the carboxy-terminal p532 RCC1-like domain and Clathrin heavy chain (CHC) respectively. The resulting transformants functioned as a positive control for a p532 (RLD2)-CHC interaction. RCC1 was amplified by PCR from a human testis cDNA library using the RCC1ATGNcoI and RCC1StopBamHI primers (see Materials and Methods, page 59) and cloned into the pACTII vector. It was transfected into Y190 with Ran (in the pACTII vector) and served as a positive control for an interaction involving Ran.

For each positive control, the yeast transformants were plated onto synthetic complete media lacking leucine, tryptophan and histidine and *HIS3*$^{+}$ colonies were assayed for β-galactosidase activity.

3.4.2 - Negative controls

Aliquots of competent Y190 yeast cells were separately transfected with each of the bait and prey constructs used in these experiments (pAS1.RPGR$^{\text{N-terminal domain}}$, pGBT9.p532$^{\text{RLD2}}$, pAS1.SNF1, pAS1.RCC1, pACTII.Ran, pACTII.ARF1, pGAD-GH.Clathrin heavy chain, pACTII.PDEδ or pACTII.SNF4) and also with empty two-hybrid vectors (pAS1 and pACTII). The yeast were spread onto media selective for the particular plasmid (i.e. synthetic complete media lacking tryptophan or leucine) and for *HIS3* expression (therefore also lacking histidine). This was carried out to determine whether any of the plasmids in isolation activated the Y190 reporter genes and provided a series of negative controls for each of the yeast two-hybrid constructs and for the vector backbones.

3.4.3 - Results

None of the candidate proteins (Ran, ARF1, Clathrin heavy chain) were found to interact with human RPGR in the yeast two-hybrid system. Only three of the four positive controls gave positive results: SNF1/SNF4, RPGR$^{\text{N-terminal domain}}$/PDEδ and p532$^{\text{RLD2}}$/Clathrin heavy chain. Yeast containing the Ran and RCC1 constructs failed to grow on the selective media and produce *HIS3*$^{+}$, *LacZ*$^{+}$ colonies. Each of the negative control transformations described in section 3.4.2 produced negative results (with no viable yeast cells growing on the selective media). These results are summarised in Table 3.1.

Table 3.1: Results of screening candidate RPGR-interactor proteins against the RPGR$^{N\text{-terminal domain}}$ using the yeast two-hybrid system.

	Bait	Prey	Interaction
Negative controls	pAS1	None	Negative
	None	pACTII	Negative
	pAS1.RPGR$^{N\text{-terminal domain}}$	None	Negative
	pGBT9.p532^{RLD2}	None	Negative
	None	pACTII.Ran	Negative
	None	pACTII.ARF1	Negative
	None	pGAD-GH.Clathrin heavy chain	Negative
	None	pACTII.PDEδ	Negative
	pAS1.SNF1	None	Negative
	None	pACTII.SNF4	Negative
Experimental	pAS1.RPGR$^{N\text{-terminal domain}}$	pACTII.Ran	Negative
	pAS1.RPGR$^{N\text{-terminal domain}}$	pACTII.ARF1	Negative
	pAS1.RPGR$^{N\text{-terminal domain}}$	pGAD-GH.Clathrin heavy chain	Negative
Positive controls	pAS1.RPGR$^{N\text{-terminal domain}}$	pACTII.PDEδ	Positive
	pGBT9.p532^{RLD2}	pGAD-GH.Clathrin heavy chain	Positive
	pAS1.RCC1	PACTII.Ran	Negative
	pAS1.SNF1	pACTII.SNF4	Positive

3.5 - SCREENING A HUMAN PANCREAS LIBRARY

At the start of this project, yeast two-hybrid libraries of retinal origin were unavailable. However, *RPGR* is ubiquitously expressed and a northern blot of human tissues probed with a radiolabelled 3' RPGR cDNA fragment indicated strong expression in the pancreas, and undetectable expression in retina, suggesting that interacting partners could be identified in this tissue. For this reason, a human pancreas cDNA library, cloned into the GAL4 vector pGAD10, was purchased from Clontech Laboratories Inc. The pGAD10 vector is similar to the pACTII vector and is compatible with the pAS1 vector with which it was used in the screens. The two vectors differ in the level of expression of the fusion protein: pACTII directs high levels of expression, whereas pGAD10 directs relatively low levels of expression. The human pancreas library contains 2×10^6 independent clones. The library culture supplied by the manufacturer was amplified and used to prepare DNA.

The RPGR RCC1-homologous domain in the pAS1 vector was used to screen the library. Approximately 3 x 10^6 clones were screened by transfecting the library and the RPGR bait into the Y190 yeast strain and plating onto synthetic complete medium lacking histidine, tryptophan and leucine but containing 25 mM 3-aminotriazole. From these yeast transformation plates, 115 putative positive colonies were identified on the basis of their histidine prototrophy. These colonies were tested for *LacZ* expression using the β-galacotosidase assay, which reduced the number of putative positive colonies from 115 to 13 (*HIS3*$^+$, *LacZ*$^+$). Library plasmids were recovered from the yeast colonies and re-tested for an interaction with RPGR by transfecting them individually into Y190 along with the RPGR bait. This time none of the 13 clones displayed histidine prototrophy or β-galactosidase activity, suggesting that they were all false-positives. Four of the 13 plasmids (31%) were subsequently found to lack cDNA inserts. The inserts from the remaining nine plasmids were sequenced and used to search the databases for identical clones. They were found to correspond to human pancreatic trypsin (accession number M22612), human dipeptidase (D13138), human chymotrypsinogen (M24400), human islet of Langerhans regenerating protein (M18963), human glypican 3 (AF003529), an EST similar to ribosomal protein S3a (M84711), two chromosome 7 sequence tagged sites (STS) (G31665 and G20212) and an uncharacterised EST (AA630545). This unacceptably high rate of false positives suggested that a change of cDNA library was required.

3.6 - SCREENING A BOVINE RETINA LIBRARY.

A second yeast two-hybrid library was obtained from Dr W. Baehr (Moran Eye Center, University of Utah Health Center, Salt Lake City). This library had been constructed by Clontech in the pGAD10 vector using bovine retina RNA. The library was amplified and used to prepare DNA which was then screened in the same way as the human pancreas library. Approximately 1.28 x 10^6 cDNA clones were screened with human RPGR in pAS1. 300 colonies were selected from the transformation plates on the basis of their histidine prototrophy and 72 of these putative positives subsequently gave positive results in the β-galactosidase assay. The library plasmids were isolated from the 72 colonies and re-tested in individual transfections. Each of the 72 gave a very weak positive result (i.e. few *HIS3*$^+$ colonies and a faint, slow-developing blue colour upon testing for β-galactosidase activity, relative to the SNF1/SNF4 positive control). 14 of the 72 plasmids were found to have no insert (19%). Common sequences were sought among the remaining 58 plasmids by PCR-amplifying the cDNA inserts (using the pGAD10F and pGAD10R primers) and then cutting

the products with the restriction endonuclease *AluI* (which is expected to cut human DNA once every 0.3 kb (Sambrook, Fritsch and Maniatis, 1989)). None of the 58 plasmid inserts gave the same restriction pattern, suggesting that they represented different clones. Firstly, it is unlikely that RPGR interacts with 58 different proteins (although it is possible that they might be different parts of the same protein). Secondly, the high rate of empty vectors weakened the argument that they were true positives. Finally, both the empty vectors and those containing inserts only showed a weak blue colour in the β-galactosidase assay, so that there was no qualitative difference between the vectors with and without inserts. No further work was done on these clones since they also appeared to be false-positives.

3.7 - SCREENING A SECOND BOVINE RETINA LIBRARY – ISOLATION OF RPGRIP

A second bovine retina yeast two-hybrid library was obtained from Dr C. Sung (Department of Ophthalmology, Cornell University Medical College). This library contained 2 x 10^6 independent clones in the pACTII vector and had been used successfully by others (Kumar *et al.*, 1996).

The library was electroporated into DH10B cells and the transformants amplified and used to prepare DNA. It was screened using the same human RPGR bait as before and an estimated 2 x 10^6 clones were screened, as calculated from the transformation control plates. 32 colonies were selected from the SC-leu/-trp/-his 3-AT medium on the basis of their histidine prototrophy. Ten of the 32 colonies gave a strong positive result in the β-galactosidase assay (a vivid, rapidly-developing blue colour - comparable to the SNF1/SNF4 positive control). Library plasmids were isolated from 8 of the 10 yeast colonies (it was not possible to recover plasmids from the other 2) and re-tested for their ability to interact with RPGR. All of the 8 colonies gave positive results although one was much weaker than the rest (fewer *HIS3*$^+$ colonies, and a faint blue colour). The inserts from the 7 plasmids that produced strong positive results were partially sequenced using the pACTinsF primer. One colony was found to contain multiple plasmids; the other 6 contained single plasmids each with a cDNA insert corresponding to the same transcript. The insert sizes were estimated by PCR-amplifying the inserts (using the pACTinsF and pACTinsREV3 primers) and comparing the products to DNA size markers. Five of the inserts were found to be approximately 760 bp in length and one was approximately 740 bp. One of the longer clones was sequenced fully and revealed a 782 bp insert consisting of a 678 bp open reading frame followed by a 78 bp polyadenine tract 3' to the in-frame stop codon. A potential

polyadenylation tract (TATAAA (Graber *et al.*, 1999)) was found 18 bp upstream of the polyadenine tract. This is shown in Figure 3.4.

```
  1  GAGATGACATCATCCTATTCAGCACAGATACTAAAGGAAACCCCACATCCTGTGAATGAT
     E  M  T  S  S  Y  S  A  Q  I  L  K  E  T  P  H  P  V  N  D

 61  AAAGAATTCTGTGAACAAGCTTCTGAAGGCAGTGAAGCGCAAACGACAGACAGCGATGAG
     K  E  F  C  E  Q  A  S  E  G  S  E  A  Q  T  T  D  S  D  E

121  ATAGTNACACCTGTNTCTCAAAAATGCCCTAAGGCAGATTCAGAGAAGATGTGCATTGAA
     I  V  T  P  V  S  Q  K  C  P  K  A  D  S  E  K  M  C  I  E

181  ATCGTCTCCCTGGCCTTCTACCCAGAGGCTGAAGTGATGTGTGACGAAAACGTGGAACAG
     I  V  S  L  A  F  Y  P  E  A  E  V  M  C  D  E  N  V  E  Q

241  GTGTATGTGGAATACAGGTTCTACGATCTGCCCTTGTCAGAGACGGAGACTCCAGTATCC
     V  Y  V  E  Y  R  F  Y  D  L  P  L  S  E  T  E  T  P  V  S

301  CTGAGGAAACCCAGAGCAGGAGAAGAAATCTACTTCCACTTCAGCAAGGTGATAGATCTA
     L  R  K  P  R  A  G  E  E  I  Y  F  H  F  S  K  V  I  D  L

361  GACCCACTGGAGCAAAAGAATCGGAGGCAGTTTCTGTTCACCATGCTGATTGGAGAAGAT
     D  P  L  E  Q  K  N  R  R  Q  F  L  F  T  M  L  I  G  E  D

421  CCTGAGCAAGGACATTTAAAGTTTACAGTAGTAAGTGATCCTATTGAGGAGGAAAAGAAA
     P  E  Q  G  H  L  K  F  T  V  V  S  D  P  I  E  E  E  K  K

481  GAATGTCAAGAAATAGGCTACGCATATCTCGAGCTGTGGCCGATGCTGGTATCAGGAAGA
     E  C  Q  E  I  G  Y  A  Y  L  E  L  W  P  M  L  V  S  G  R

541  GACATCCTAGAGCAAGATCTAGACATTGTTGGCCCTGAAGATCAGGCTACCCCCATAGGA
     D  I  L  E  Q  D  L  D  I  V  G  P  E  D  Q  A  T  P  I  G

601  AAGCTCAAGGTGTCTCTTCAAGCTGCTGCTGCCCTCCAAGCTATTTACAAGGAGATGACT
     K  L  K  V  S  L  Q  A  A  A  A  L  Q  A  I  Y  K  E  M  T

661  GAAGATTTGTGTTCCTGATGGAACAAGTGCTAGTCCATTCTAAACCTGAGGGAACCATAG
     E  D  L  C  S  *  W  N  K  C  *  S  I  L  N  L  R  E  P  *

721  TCAAAAGTCTCTTATAAAGTAACTTGCTCTACTATGAAAAAAAAAAAAAAAAAAAAAAAA
     S  K  V  S  Y  K  V  T  C  S  T  M  K  K  K  K  K  K  K  K

781  A
```

Figure 3.4: Sequence of the RPGR-interacting clone isolated from a bovine retina yeast two-hybrid library. Translation of the open reading frame is shown and the in-frame stop codon shown in bold. A potential polyadenylation signal is underlined.

One of the positive clones was tested for the ability to activate the Y190 reporter genes in the absence of a second (bait) plasmid and also for the ability to interact with RCC1, Ran and SNF1. The results of all these tests were negative, suggesting that the original positive result was dependent on and specific to RPGR.

A search of the EST databases for sequences similar to the bovine interactor identified 10 human (r93221, ai964059, ai655818, ai632512, ai05922, aa782963, aw681763, aa928161, aa476670 and w26173) and 8 mouse (aa204546, c80430, aa154084, ai451895,

av258740, aw050180, aa208719, av155762) sequences, all previously uncharacterised, containing 86% (human) and 76% (mouse) nucleotide identity to the bovine sequence. The IMAGE EST clone r93221, which showed homology to the 5' end of the bovine interactor, was obtained from the MRC HGMP resource centre (MRC HGMP RC). The cDNA insert of this clone was sequenced, which revealed homology to the bovine clone extending as far as the stop codon. The mouse ESTs corresponded only to the 3' end of the bovine and human sequences. Figure 3.5 shows translations of the human, bovine and mouse (partial) sequences in alignment.

```
human     MTLSHSALKQKEPLHPVNDKESSEQGSEVSEAQTTDSDDVIVPPMSQKYPKADSEKMCIE
bovine    MTSSYSAQILKETPHPVNDKEFCEQASEGSEAQTTDSD-EIVTPVSQKCPKADSEKMCIE
mouse     ------------------------------------------------------------

human     IVSLAFYPEAEVMSDENIKQVYVEYKFYDLPLSETETPVSLRKPRAGEEIHFHFSKVIDL
bovine    IVSLAFYPEAEVMCDENVEQVYVEYRFYDLPLSETETPVSLRKPRAGEEIYFHFSKVIDL
mouse     -----FCPEADVHVLXDYPQVYVEYKFCESPLSETETPMSLRRPRAGEEIHFHFSKVIDL
               * ***:*    :  ******:* : ********:***:*******:*********

human     DPQEQQGRRRFLFDMLNGQDPDQGHLKFTVVSDPLDEEKKECEEVGYAYLQLWQILESGR
bovine    DPLEQKNRRQFLFTMLIGEDPEQGHLKFTVVSDPIEEEKKECQEIGYAYLELWPMLVSGR
mouse     DPVEHQSRTVHLFAMLHAQDSDEGRFKFTVVSDPLDEEKKECQDIGYAYLELWQIFQSGK
          ** *:::.*  .** ** .:*.:::*::********::******:::*****:** :: **:

human     DILEQELDIVSPEDLATPIGRLKVSLQAAAVLHAIYKEMTEDLFS
bovine    DILEQDLDIVGPEDQATPIGKLKVSLQAAAALQAIYKEMTEDLCS
mouse     DILEQELEIVSPRNQAIQIGRLKVSLQAAAALHGIYKEMTEDLFS
          *****:*:**.*.: *  **:*********.*:.********* *
```

Figure 3.5: CLUSTAL W alignment (Thompson *et al.*, 1994) of the translations of the human, bovine and (partial) mouse RPGR-interactor sequences (* = single, fully conserved residue, : = conservation of strong groups, . = conservation of weak groups, blank space = no consensus).

The open reading frame of the r93221 human EST, from the first ATG triplet to the stop codon, was PCR-amplified using primers r93221NcoI and r93221BamHI and subcloned into the pACTII vector. This human construct was treated as a candidate RPGR-interactor using the yeast two-hybrid system as described above. It was found to interact with human RPGR but not with RCC1, Ran or SNF1, nor did it activate the Y190 reporter genes in the absence of a bait construct (Figure 3.6).

This new RPGR-interacting protein was given the name RPGRIP (RPGR-Interacting Protein). Human RPGRIP (hRPGRIP) was tested for the ability to interact with the C-terminal portion of RPGR (encoded by exons 13 to 19), PDEδ and the small GTPase Rab13. It was of interest to determine whether RPGRIP interacted with PDEδ because this protein is known to interact with RPGR (Linari *et al.*, 1999). Rab13 was tested because this had

previously been shown to interact with PDEδ (Marzesco *et al.*, 1998). RPGR exons 13 to 19 were PCR amplified as a single fragment from an RPGR cDNA clone derived from an adult brain cDNA library using the Exon13NcoI and StopBamHI primers. The fragment was subcloned into pAS1. Rab13 cDNA was PCR amplified from the pGAD10 human pancreas library using primers Rab13NcoI and Rab13BamHI and also subcloned into pAS1.

Human RPGRIP did not activate the Y190 reporter genes in the absence of a second plasmid. RPGRIP did not interact with the C-terminal RPGR construct, PDEδ or Rab13.

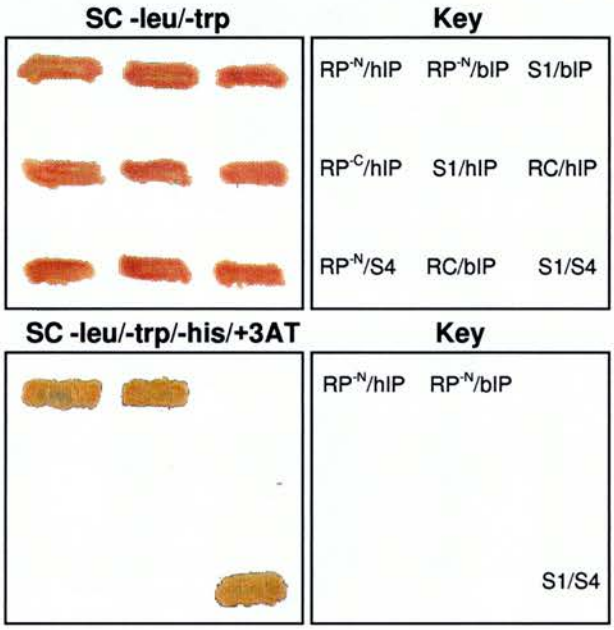


Figure 3.6: A novel protein interacts with the RCC1-like domain of RPGR ($RPGR^{-N}$) in the yeast two-hybrid system but does not interact with the C-terminal half of RPGR, RCC1 or an unrelated protein (SNF1). Y190 yeast transformed with bait and prey constructs were plated onto SC –leu/-trp to confirm transformation and SC-leu/-trp/-his/+3AT to test for interacting proteins (pAS1construct/pACTIIconstruct: hIP, human RPGRIP: bIP, bovine RPGRIP: $RPGR^{-C}$, RPGR C-terminal domain: RC, RCC1: S1, SNF1: S4, SNF4).

3.8 - TESTING RPGRIP FOR INTERACTION WITH MUTANT RPGR CONSTRUCTS.

An experiment was carried out to determine whether the interaction between hRPGRIP and RPGR is disrupted when point mutations associated with xlRP are present in the *RPGR* coding sequence. Nine *RPGR* constructs were obtained from Dr. J. Becker (Max Planck Institut für Molekulare Physiologie, Dortmund) in the pBTM116 two hybrid vector (Linari *et al.*, 1999). The plasmids contained RPGR cDNA from the start of exon 1 to the middle of exon 10, encoding amino acids 1-392 (the RCC1-like domain of RPGR encompasses amino

acids 39 to 365 ((Renault *et al.*, 1998)). One of the constructs contained the wild-type sequence, seven contained XLRP disease-associated *RPGR* mutations (G60V, H98Q, F130C, G215V, P235S, C250R, and G275S) (Meindl *et al.*, 1996; Roepman *et al.*, 1996) and one contained a substitution that was found in a patient with X-linked congenital stationary night blindness (V36F; T. Meitinger, personal communication). The inserts were cut out of pBTM116 (which is used in lexA-based yeast two-hybrid systems) as *Bam*HI/*Pst*I fragments and subcloned into pAS1. The wild-type construct was also tested for an interaction with hRPGRIP because the RPGR bait used in the library screen coded for the first 502 amino acids whereas the constructs obtained from J. Becker coded for the first 392 amino acids.

The wild-type RPGR was found to interact with RPGRIP in the yeast two-hybrid system but the mutants either showed absent (V36F, G60V, H98Q, F130C, P235S, C250R) or slightly decreased (G215V, G275S) reporter gene activity (Figure 3.7). These results are summarised in Table 3.2. All of the disease-causing mutations tested, with the exception of F130C, alter residues of the RPGR RCC1-like domain that are conserved in vertebrates (Renault *et al*, 1998), and are therefore likely to be important for correct folding and function (shown in Figure 3.8).

Figure 3.7: Y190 yeast transformed with wild-type or mutant RPGR bait (in pAS1) and RPGRIP prey (in pACTII) constructs were plated onto SC –leu/-trp to confirm transformation and SC-leu/-trp/-his/+3AT to test for interacting proteins. This showed that RPGRIP interacts with wild-type RPGR and with the G215V and G275S RPGR mutant proteins but does not interact with other mutant proteins, including V36F, G60V, H98Q, F130C, P235S or C250R RPGR. (SC-leu/-trp=synthetic complete medium lacking leucine and tryptophan, SC-leu/-trp/-his/+3AT= synthetic complete medium containing 25mM 3-aminotriazole, lacking leucine, tryptophan and histidine, SNF1/SNF4=positive control.)

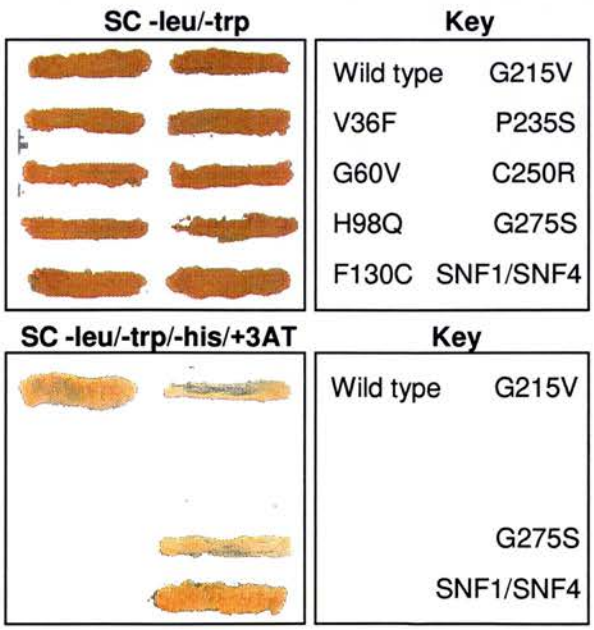


Table 3.2: Summary of hRPGRIP interaction screens using wild-type and mutant RPGR and control sequences in the yeast two-hybrid system.

	Bait	Prey	Interaction
	hRPGRIP	$RPGR^{RCC1\text{-}Like\ Domain}$	Positive
Negative controls	hRPGRIP	RCC1	Negative
	hRPGRIP	Ran	Negative
	hRPGRIP	SNF1	Negative
	hRPGRIP	none	Negative
	hRPGRIP	$RPGR^{C\text{-}Terminal\ Domain}$	Negative
	hRPGRIP	PDEδ	Negative
	hRPGRIP	Rab13	Negative
Mutants	hRPGRIP	$RPGR^{V36F}$	Negative
	hRPGRIP	$RPGR^{G60V}$	Negative
	hRPGRIP	$RPGR^{H98Q}$	Negative
	hRPGRIP	$RPGR^{F130C}$	Negative
	hRPGRIP	$RPGR^{G215V}$	Positive (weak)
	hRPGRIP	$RPGR^{P235S}$	Negative
	hRPGRIP	$RPGR^{C250R}$	Negative
	hRPGRIP	$RPGR^{G275S}$	Positive (weak)

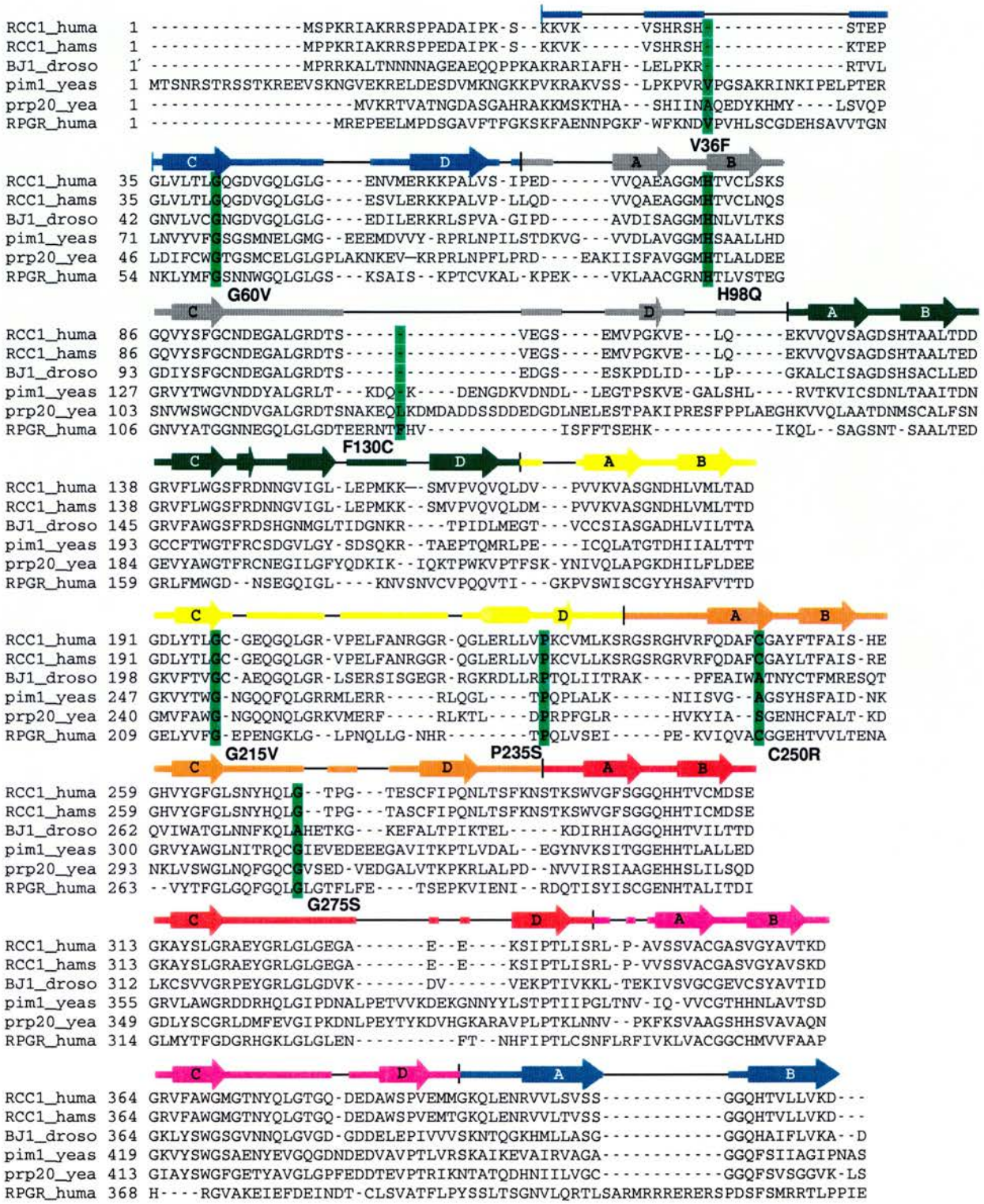


Figure 3.8: Primary structure alignment of five RCC1 orthologues and RPGR (RCC1-homologous domain only). The highlighted (green) residues show those that were mutated in human RPGR yeast two-hybrid constructs tested for an interaction with hRPGRIP. The arrows indicated the residues that form the propeller blades of the tertiary structure of human RCC1 (Renault *et al.*, 1998).

3.9 - CO-IMMUNOPRECIPITATION OF RPGR AND RPGRIP

The proposed interaction between RPGR and RPGRIP was tested by co-immunoprecipitation of *in vitro* translated RPGR and RPGRIP proteins. Oligonucleotide primers were used to amplify the cDNA inserts from pAS1.RPGR and pACTII.RPGRIP. The PCR primers are designed so that the translated products contain epitope tags at one end. RPGR was tagged with c-myc and RPGRIP with haemagglutinin (HA) epitopes. The amplified sequences were used in a coupled *in vitro* transcription/translation reaction to produce [^{35}S] radiolabelled proteins. The proteins were mixed and immunoprecipitated with either anti-myc or anti-HA antibody as detailed in Chapter 2 (page 80). Control reactions were performed by mixing RPGR-myc with HA-tagged and radiolabelled Max protein or RPGRIP-HA with myc-tagged and radiolabelled lamin protein and immunoprecipitating with anti-myc and anti-HA respectively. The results of this experiment are shown in Figure 3.9. The four lanes on the right hand side of the autoradiograph show undiluted *in vitro* transcription/translation products separated on the same gel to allow identification of the bands in the four test lanes (shown on the left of the figure). From a mixture of RPGR-myc and RPGRIP-HA it was possible to co-immunoprecipitate both proteins using either of the two antibodies. From a mixture of RPGR-myc and Max-HA only RPGR-myc was immunoprecipitated with anti-myc. From a mixture of RPGRIP-HA and lamin-myc only RPGRIP-HA was immunoprecipitated with anti-HA. This suggests that the interaction between RPGR and RPGRIP is specific and not an artefact of the yeast two-hybrid system. The bands obtained with anti-*c*-myc mouse monoclonal antibody (Figure 3.9, lanes 1, 3) are weaker than those obtained with anti-HA rabbit polyclonal antibody (lanes 2, 4), either because protein G, which was used to capture the antigen-antibody complexes, binds more strongly to rabbit than to mouse antibodies, or because the polyclonal antibody binds with higher avidity than the monoclonal (Harlow and Lane, 1988).

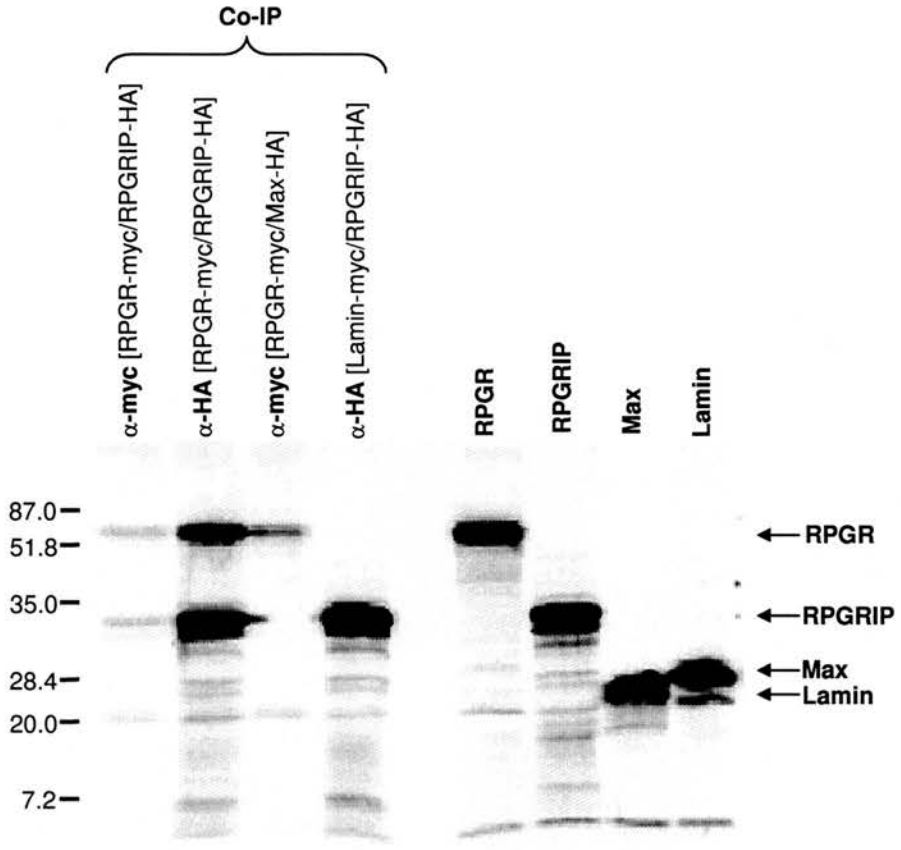


Figure 3.9: Co-immunoprecipitation of RPGR and RPGRIP *in vitro* transcribed/translated proteins. Lane 1, RPGR-myc plus RPGRIP-HA immunoprecipitated with anti-myc antibody; lane 2, RPGR-myc plus RPGRIP-HA immunoprecipitated with anti-HA antibody; lane 3, RPGR-myc plus Max-HA immunoprecipitated with anti-myc antibody; lane 4, RPGRIP-HA plus lamin-myc immunoprecipitated with anti-HA antibody. The four lanes on the right hand side of the gel contain undiluted *in vitro* transcription/translation products for size comparison with proteins on the left hand side. The position of molecular weight markers are indicated on the left of the gel (sizes in kDa).

3.10 - DISCUSSION

The yeast two-hybrid system is a widely used method for detecting interacting proteins and for gaining insights into protein function and pathways (Luban and Goff, 1995). This is an *in vivo* method for detecting protein-protein interactions and requires confirmation with one or more independent methods, but provides a powerful means of screening large numbers of cDNA clones for interacting products (Fields and Song, 1989; Chien *et al.*, 1991).

The yeast two-hybrid system has a number of advantages over other methods of detecting interacting proteins. Firstly, the experiments are carried out *in vivo*. This increases the likelihood that the proteins under investigation will adopt their native conformations. In

addition, unlike bacterial systems, some post-translational modifications are made to proteins expressed in yeast. The yeast two-hybrid system is capable of observing relatively weak interactions. For example, the interaction between the mammalian RAF protein kinase with the Ras GTPase protein was detected in the yeast two-hybrid system but not by co-immunoprecipitation methods (van Aelst *et al.*, 1993). The source of this sensitivity is, in part, the three-fold amplification of positive signals, i.e. transcriptional, translational, and enzymatic. It is likely that even a transient interaction can activate transcription to produce a stable mRNA, which can then be repeatedly translated (Fields and Sternglanz 1994). Another valuable feature of the system is the presence of the *HIS3* reporter gene, encoding histidine prototrophy, which enables high density plating and stringent selection of yeast transformants expressing interacting proteins. This allows greater numbers of library clones to be screened than if colour selection (e.g. *LacZ*) was alone available. One other advantage of the yeast two-hybrid system over other methods of purifying interacting proteins is that the end product of a library screen is a cDNA clone of the interactor, rather than a polypeptide. Lastly, by fragmenting the interacting cDNA clones it is possible to determine the limits of the interaction domains.

Not all protein-protein interactions can be detected in the yeast two-hybrid system. Some proteins are unsuitable as baits because they can activate the reporter genes in the absence of an interacting partner. In these cases, it is necessary to delete whichever portions of the protein are responsible for this activity – if this is possible (Bartel *et al.*, 1993b). In some cases, the presence of the DNA-binding domain or activation domain may interfere with the correct folding of the fusion protein or occlude the normal site of interaction and so impair the ability to interact (van Aelst *et al.*, 1993). In this situation it may help to switch the transcription factor domain from the N-terminal end (the usual position) to the C-terminal end of the protein under investigation (Béranger *et al.*, 1997). The conditions within the yeast cell may preclude correct folding or post-translational modification of some mammalian proteins. In these cases, interactor hunts might fail as a result, conversely some interactions that are observed in yeast may not occur in the protein's native environment (Fields and Sternglanz 1994). Similarly, (in the yeast two-hybrid system) it may not be possible to detect interactions involving extracellular proteins – these often contain disulphide bonds and/or are glycosylated, and these modifications are not generally compatible with a nuclear-based system (Fields And Sternglanz 1994). Despite these limitations, mammalian proteins produced in yeast are more likely to resemble their native state than those produced in bacterial expression systems (Fields and Sternglanz, 1994).

The function of the RPGR protein is unknown, although Linari *et al.* (1999) previously showed that it interacts with PDEδ by screening an embryonic mouse cDNA library using a yeast two-hybrid system. This interaction was confirmed by *in vitro* protein binding (pull-down) assays and affinity measurements (K_D = 90 nM), which were consistent with a physiological interaction. In the present study, a second RPGR-interacting protein has been identified, RPGRIP, by screening a bovine retinal cDNA library with the human RPGR N-terminal domain in a yeast two-hybrid system. Six out of seven strongly positive clones were shown to contain fragments of the bovine *RPGRIP* gene and the interaction was confirmed with human RPGRIP. The bait construct contained the first 12 exons of RPGR, including the whole of the RCC1-like domain. This construct was chosen because the goal of the experiment was to identify proteins that interact with the RCC1-like domain, which, at the time this work was carried out, contained all of the known RP3 mutations (Meindl *et al.*, 1996). An additional line of evidence for the interaction between RPGR and RPGRIP was provided by the demonstration of specific co-immunoprecipitation of the complex following *in vitro* transcription and translation of the two gene products.

Interpreting the results of the candidate interacting protein screens is problematic because of the failure of Ran and RCC1 to interact in this system, which was unexpected. It is not always possible to have positive controls for all constructs. For example, before the interaction between RPGR and PDEδ was detected there was no positive control for RPGR. If the Ran-RCC1 control had not been included then it would have been concluded that RPGR does not interact with Ran. Similarly, the conclusion that ARF1 does not interact with RPGR has to be guarded in the absence of a positive control for ARF1. The negative results obtained for ARF1 and Ran are supported by the absence of detectable GEF activity for RPGR in the presence of Ran or ARF1 (F.Manson, unpublished results). Linari *et al.* (1999) reported a similar result for Ran, which also tends to support the results from these experiments.

Why did Ran and RCC1 not interact in this system? Both proteins are present in the nucleus in humans. It may be that the presence of the GAL4 domain in the RCC1 hybrid prevented the RCC1 protein from adopting its complex seven-bladed propeller structure (Renault *et al.*, 1998). The C-terminus and a region close to the N-terminus of RCC1 come together to form the first blade of the propeller. The GAL4 domain is attached to the N-terminus of RCC1 and it is possible that this prevents correct structural organisation of RCC1. Attachment of the GAL4 domain to the C-terminus would be even more likely to disrupt the three-dimensional structure, as this is directly involved in forming the first structural repeat of RCC1. Ran and RCC1 are known to interact: they were isolated as a 1:1

complex from HeLa cell nuclei upon chelation of Mg^{2+} ions (Bischoff *et al.*, 1990). This complex is stable in the presence of 2 mM Mg^{2+} and the absence of GTP or GDP but completely dissociates in the presence of GTP or GDP and Mg^{2+} (Bischoff and Ponstingl 1991). Of course, RCC1 must interact with Ran in the cell in order to stimulate guanine nucleotide dissociation and Ran has been observed in a ternary complex with RCC1 and guanine nucleotide, indicating that these two factors do not compete for the same binding site on Ran (Klebe *et al.*, 1995). Klebe *et al.* (1995) proposed a scheme for guanine nucleotide exchange whereby RCC1 binds to Ran-GDP, GDP dissociates (leaving the binary Ran-RCC1 complex), GTP binds to Ran-RCC1 and RCC1 dissociates from Ran-GTP. In this scheme the Ran-RCC1 interaction would, therefore, be a transient one and possibly too short-lived to allow detection in the yeast two hybrid system. Klebe *et al.* (1995) also reported the affinities of the Ran-GTP/RCC1 and Ran-GDP/RCC1 interactions as being relatively low, at 18 μM and 13 μM respectively, compared with the 90 nM RPGR-PDEδ interaction.

It is also possible that over-expression of Ran and RCC1 is harmful to yeast cells. In mammalian cells, over-expression of Ran prolongs the duration of S-phase and leads to reduced cell proliferation and decreased total cell numbers (Milano and Strayer, 1998). Over-expression of the yeast RCC1 homologue *pim1* causes cells to arrest with postmitotically condensed chromosomes, an unreplicated genome and a wide medial septum (Matynia *et al.*, 1998). However, previous examples have shown both Ran and other exchange factors can be overexpressed without problems in the yeast two-hybrid system. For example, Schürmann *et al.* (1999) used ADP-ribosylation factor-related protein (ARP) as bait and identified the ARF-specific guanine nucleotide exchange factor Cytohesin as an interacting partner in a two-hybrid library screen. Ran was used as bait to identify the Xenopus Ran-binding protein 1 (Nicolas *et al.*, 1997), and, in *Arabidopsis*, a mutant form of Ran, which remains in the GTP-bound form, was used to identify the *Arabidopsis* Ran-binding protein 1. A number of human Ran mutants are known to be locked in the GTP-bound form due to an absence of GTPase activity including Ran G19V (Bischoff *et al.*, 1994), Ran Q69L and Ran L45E (Lounsbury *et al.*, 1996). The T24N Ran mutant is known to have an approximately 100-fold higher affinity for RCC1 than wild-type Ran (Lounsbury *et al.*, 1996). Although this would have been a preferable alternative to wild-type Ran in this test, the wild-type sequence was appropriate in the first instance. Time constraints prevented the construction of such mutants for screening in the yeast two-hybrid system.

p532 was initially localised to the Golgi complex by Rosa *et al.* (1996). This localisation, and the presence of two RCC1-like domains in the p532 sequence, led the

authors to investigate p532 binding to small GTPases associated with the Golgi complex. ARF1, Rab3a, and Rab5 were all tested using co-immunoprecipitation methods but only ARF1 bound to p532-RLD2. It was for this reason that ARF1 was considered a candidate RPGR-interactor. RLD-1 was found to act as a GEF for the ARF1, Rab 3A and Rab5 proteins. To only ARF1, however, was p532 able to both bind (RLD2) and show GEF activity (RLD1), suggesting a physiological role.

Screening two different two-hybrid libraries in pGAD10 yielded no confirmed positive clones and a large number of false positive results. The first screen identified 13 putative positive clones (*HIS3*$^+$, *LacZ*$^+$) all of which turned out to be false positives. One-third of these clones lacked cDNA inserts, while the remainder contained a variety of clones that would be expected to be abundant in a pancreatic library and therefore represent false positives.

The identification of false positives is a known hazard of the yeast two-hybrid system. These are frequently the result of a library plasmid encoding a protein involved in transcription. A list of yeast two-hybrid false-positives reported by various workers can be found at the web site *http://www.fccc.edu/research/labs/golemis/intro.html*. Screening the bovine retina library in pGAD10 also failed to produce any true positives. This time the number of putative positive clones as determined by histidine prototrophy was larger (300), and the number displaying both *HIS3* and *LacZ* expression was much larger (72). However, all of these clones again appeared to be false positives when treated as candidate interactors and 14 of them (19%) lacked inserts. One possibility is that some feature of the pGAD10 vector backbone is responsible for inappropriate activation of the reporter genes. The product of pGAD10 consists of the insert protein fused to amino acids 768 to 881 of GAL4, the same as in pACTII. According to the vector manufacturers, pGAD10 is perfectly compatible with pAS1 and the only significant difference between pGAD10 and pACTII is that the latter drives much higher expression of the fusion protein in yeast. An excess of activation domain hybrids over DNA-binding domain hybrids is desirable. This ensures that more of the DNA-binding domain hybrids that are occupying UASs are likely to associate with activation domain hybrids and activate reporter gene expression (Fields and Sternglanz, 1994). If the lower expression of the pGAD10 hybrids leads to an excess of DNA-binding domain plasmids then a lowering of reporter gene expression is expected, but this does not explain the large number of false positives observed.

It is not clear why PDEδ was not detected on two-hybrid screening of the bovine retina library, since it is known to be strongly expressed in this tissue (Florio, Prusti, and Beavo, 1996) and the interaction has been repeatedly confirmed by direct screening. The

absence of PDEδ from the positive clones described in section 3.7 can be compared with the results of other groups attempting to isolate RPGR-interacting proteins. Eight out of 40 clones isolated by Linari *et al.* (1999) in a yeast two-hybrid screen of a mouse embryo cDNA library encoded the PDE delta subunit. The authors do not state how many clones were screened in total, or how many the library contained. Hong *et al.* (2000) screened 1 x 10^6 clones from a mouse retina library containing 1 x 10^6 clones in total. Out of six positive clones, one encoded PDEδ and five encoded mRPGRIP. Roepman *et al.* (2000) screened 6 x 10^6 bovine retina clones in the yeast two-hybrid system in an attempt to identify RPGR-interacting proteins and isolated six *RPGRIP* clones but no PDEδ clones.

Possible reasons why PDEδ was not detected in this study include differences in the RPGR "bait", a species difference, and the exhaustiveness of library screening. Linari *et al.* (1999) used an *RPGR* bait containing amino acids 1-392, whereas here a bait consisting of amino acids 1-502 was used (the RLD consists of residues 39-365). It seems unlikely that this difference would account for the failure to identify PDEδ. A second possibility is a species difference, since the RPGR-PDEδ interaction was detected by screening a mouse library with human RPGR, whereas a bovine library was used in this study. This is also an unlikely explanation however, since human, mouse and bovine PDEδ are unusually well conserved, with 98% of residues identical at the amino acid level between the three species. Only 3 conservative substitutions out of 150 residues separate these species and only 1 substitution occurs between bovine and human (compared with 2 between bovine and mouse; see Figure 1.15, page 46). This leaves the third possibility, namely a failure to exhaustively screen the library. An estimated 2 x 10^6 clones were screened and prior to amplification the library contained 2 x 10^6 independent clones (Kumar *et al.*, 1996). Amplifiying the library may have led to the underrepresentation of some clones. The probability of detecting a specific clone in a cDNA library can be estimated using the method described in Hartl and Jones (1988). The number of clones required to be screened (n) in order to achieve a probability (P) of detecting a specific clone (which accounts for a fraction (f) of the total number of clones in the library) is given by:

$$n = \ln (1-P) / \ln (1-f)$$

Thus, if there was a single PDEδ clone in the bovine retina library, and assuming the library contained 2 x 10^6 clones in total, the number of clones required to be screened in order to obtain a 90% probability of detecting the PDEδ clone is:

$$\ln (1\text{-}0.9) / \ln (1\text{-}5 \times 10^{-7}) = 4{,}605{,}169$$

If the fraction of PDEδ clones is increased to 10 in 2 x 10^6 clones, the number of clones that have to be screened to detect PDEδ is reduced accordingly (to ln (1-0.9) ÷ ln (1-5 x 10^{-6}) = 460,516). In the same way, to increase the probability of finding a clone larger numbers of clones must be screened. Table 3.3 shows how the probability of detecting PDEδ by screening 2 x 10^6 clones changes according to the proportion of PDEδ clones in the bovine retina cDNA library.

Table 3.3: Estimates of the probability (P) of detecting PDEδ in the bovine retina yeast two-hybrid library (consisting of 2 x 10^6 unamplified clones) according to the method described in Hartl and Jones (1998). Through screening 2 x 10^6 clones, the probability of detecting PDEδ increases according to its abundance (see text).

Number of PDEδ clones in library	Probability of detection (P) by screening 2 x 10^6 clones
1 in 2 x 10^6	0.632
2 in 2 x 10^6	0.865
3 in 2 x 10^6	0.950
4 in 2 x 10^6	0.982
5 in 2 x 10^6	0.993
10 in 2 x 10^6	>0.999

Although PDEδ is reported to be abundantly expressed in the retina (Florio, Prusti, and Beavo, 1996), amplification of the library may have led to this transcript being under-represented. It remains a possibility that further screening would identify PDEδ. If this was the reason for the absence of PDEδ among the positives then it follows that there may be additional RPGR-interacting proteins in the library that were not identified for the same reason.

Further confirmation of the specificity of the RPGR-RPGRIP interaction was obtained when six out of eight XLRP disease-associated RPGR mutations (V36F, G60V, H98Q, F130C, P235S and C250R but not G215V or G275S) were shown to abolish the interaction in yeast two-hybrid experiments. These mutations all affected conserved residues and will be discussed in turn.

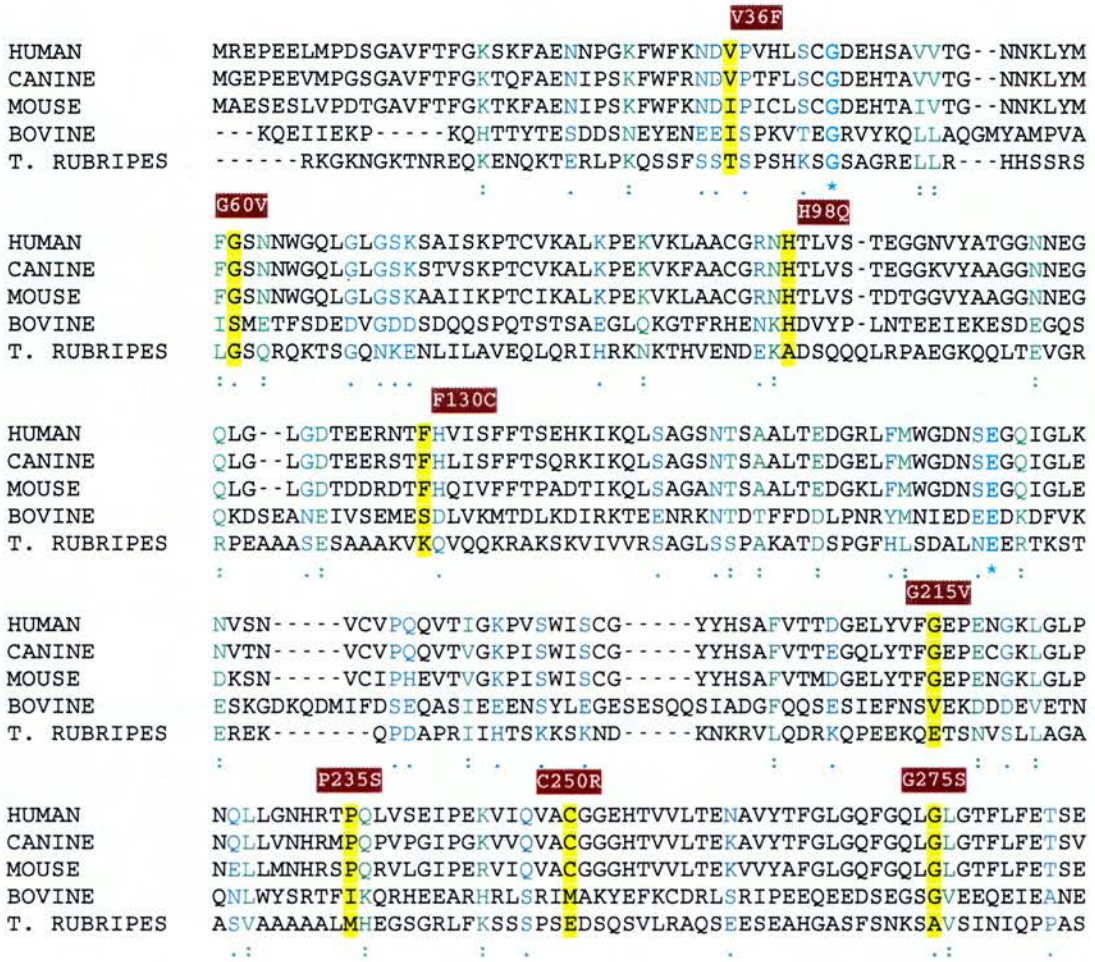


Figure 3.10: CLUSTAL W alignment (Thompson *et al.*, 1994) of human, canine, mouse, bovine and *Takifugu rubripes* RPGR polypeptide sequences from hRPGR residue 1 to 285. Positions of the substitutions investigated in the yeast two-hybrid system (section 3.8) are highlighted (yellow) (* = single, fully conserved residue, : = conservation of strong groups, . = conservation of weak groups, blank space = no consensus).

The V36F substitution was found in a single patient with X-linked congenital stationary night blindness (T. Meitinger, personal communication). On the alignment of RCC1 homologues (Figure 3.8) the valine at this position in RPGR is conserved only in the *Schizosaccharomyces pombe* pim1 protein. At this position in PRP20 (*Saccharomyces cerevisiae*) there is an alanine residue and in the other species shown in the alignment (human, hamster and *Drosophila*) there is a gap. Alanine and valine are both small amino acids with aliphatic, non-polar side chains, suggesting that they could substitute for one another in different species. An alignment of RPGR protein sequences from *Homo sapiens*, *Canis familiaris*, *Mus musculus*, *Bos taurus* and *Takifugu rubripes* is shown in Figure 3.10 and summarised (with respect to the residues at hRPGR positions 36, 60, 98, 130, 215, 235,

250 and 275) in Table 3.4. This shows that there is an aliphatic, non-polar side chain-bearing amino acid (valine or isoleucine) at this position in each of the mammalian sequences. The presence of phenylalanine at this position in RPGR abolishes the interaction with RPGRIP, which it does not for PDEδ (Linari *et al.*, 1999). This amino acid has an aromatic side chain and may not be expected to substitute for valine at a position essential for the RPGR-RPGRIP interaction. The fact that this change has no effect on the RPGR-PDEδ interaction suggests that different parts or residues of the protein may be involved in binding to PDEδ and RPGRIP. The seven remaining mutations (G60V, H98Q, F130C, G215V, P235S, C250R and G275S) were found in XLRP patients. The glycine at position 60 is conserved in the RCC1 protein of all species examined, while in RPGR the only species which deviates is *Bos taurus* (serine). The substitution of a glycine for a valine is not a radical one, suggesting that the residue is important for RPGR in terms of the interaction with RPGR. The histidine at hRPGR position 98 is 100% conserved in all of the RCC1 homologues and all of the mammalian RPGR homologues, indicating that it may be an critical residue. The replacement of the basic histidine with the amidic glutamine abolished the interaction with RPGRIP. The phenylalanine at position hRPGR 130 is conserved in canine and murine RPGR but is substituted for a serine in bovine and a lysine in *Takifugu rubripes*. The aligned RCC1 sequences (Figure 3.8) have a gap at this position, with the exception of PRP20 (*Saccharomyces cerevisiae*), which has a leucine residue. This residue is thus well conserved but restricted to RPGR and the presence of the small, sulphur-containing cysteine, in place of the aromatic phenylalanine, disrupted the interaction with RPGRIP. In RCC1, the glycine at hRPGR position 215 is 100% conserved across the species examined. The murine, human and canine RPGR proteins all have a glycine here, while there is a valine in bovine and glutamine in *Takifugu rubripes*. The substitution of a glycine for a valine failed to disrupt the interaction with RPGRIP, but did prevent binding to PDEδ (Linari *et al.*, 1999). Both of these amino acids are small, non-polar residues but the fact that the same alteration at position 60 produced a different result suggests that position 215 is less critical for the RPGRIP interaction. Changing the imino-acid proline at position 235 to a hydroxyl-containing serine also abolished the interaction with RPGRIP, although it failed to completely prevent the interaction with PDEδ (Linari *et al.*, 1999). The proline at this position is conserved in RCC1 in all species examined and in human, canine and murine RPGR. Human, canine and murine RPGR contain a cysteine at position 250 (bovine RPGR has a methionine here, and *Takifugu rubripes* a glutamate) as do all the RCC1 proteins examined except BJ1 (*Drosophila*), which has a threonine. Substituting this cysteine in RPGR for an arginine prevented binding to RPGRIP. The glycine at hRPGR position 275 is

well conserved among the RPGR homologues – all but *Takifugu rubripes* (alanine) have glycine – but less so in RCC1. RPGR and PRP20 (*Saccharomyces cerevisiae*) have the same residue here while pim1 (*Schizosaccharomyces cerevisiae*) contains an arginine, BJI (*Drosophila*) a lysine, and hamster and human RCC1 a histidine. Changing this glycine to a serine abolished the RPGR-PDEδ interaction but did not affect RPGR binding to RPGRIP.

Table 3.4: Comparison of RPGR amino acids residues at specific positions in five species. Residues were altered at these positions, as indicated in the 'Mutants' row, and tested for interaction with RPGRIP in the yeast two hybrid system. Positions refer to the human RPGR sequence. See sections 3.8 (Testing RPGRIP for interaction with mutant RPGR constructs) and 3.10 (Discussion).

Position	**36**	**60**	**98**	**130**	**215**	**235**	**250**	**275**
Mutation†	F	V	Q	C	V	S	R	S
Human RPGR	V	G	H	F	G	P	C	G
Canine RPGR	V	G	H	F	G	P	C	G
Mouse RPGR	I	G	H	F	G	P	C	G
Bovine RPGR	I	S	H	S	V	I	M	G
Takifugu rubripes*	T	G	A	K	E	M	E	A
Mutation abolishes interaction?	Yes	Yes	Yes	Yes	No	Yes	Yes	No

* retinitis pigmentosa GTPase regulator-like protein, †all except V36F were found in XLRP patients. V36F was found in a patient with CSNB.

In summary, the XLRP-associated G215V and G275S substitutions did not prevent RPGRIP binding but did prevent the interaction with PDEδ, and the V36F change, which is not associated with XLRP, disrupted RPGRIP binding only. Although these results are purely qualitative, they suggest that different RPGR sub-domains within the RLD are likely to be involved in the interactions with RPGRIP and PDEδ and that the latter may be more important for full expression of the XLRP disease. Useful future experiments would be (i) a semi-quantitative β-galactosidase assay to give a measure of the interaction between RPGR and the G215V and G275S mutants and (ii) attempts to map these sub-domains.

Following the identification of interacting partners, it is necessary to verify that binding is reproducible using independent techniques and not an artefact of the yeast two-hybrid system. The technique used here involved the *in vitro* co-immunoprecipitation of *in vitro* transcribed and translated epitope-tagged RPGR and RPGRIP. This showed that the interaction between RPGR and RPGRIP is specific and reproducible *in vitro* and strongly

supports the yeast two-hybrid *in vivo* evidence for binding. This is one of several commonly used methods to test yeast two-hybrid data. Alternatives include the co-immunoprecipitation of native proteins from tissue or cell extracts, co-immunoprecipitation of recombinant proteins from transfected cell-lines, affinity chromatographic techniques and 'pull-down' assays.

Co-immunoprecipitation of the proteins from an appropriate tissue extract (e.g. retina in the case of retinal proteins) is the ideal method because the proteins are in their normal environment and native state and a positive result strengthens the argument that the interaction occurs *in vivo*. In this, antibodies against one or both of the two interacting proteins are used to purify the complex from a tissue lysate. The antibodies are removed from the mixture using protein A or G and separated from the bound proteins, which are analysed by SDS-PAGE. The technique requires that antibodies be raised against the interacting partners and can be difficult to perform successfully if the proteins are of too low abundance for detection or antibodies are low affinity. In addition, the source tissue itself may not be easy to obtain, as is the case with human retina. A variation of this technique is to use, as a source of the binding complex, cell lines derived from the tissue in which the interaction is expected to occur. These cells can be grown in the presence of radioactive amino acids (for example [^{35}S] methionine) in order to produce radiolabelled proteins which can be more readily detected using autoradiographic methods. Alternately, the cell lines can be transfected with expression vectors encoding the interacting proteins. In this way, large amounts of the polypeptides can be obtained for co-immunoprecipitation.

Another method of verifying two-hybrid interactions involves affinity chromatography. One of the proteins is expressed containing a tag at the amino- or carboxyl-terminus (for example, a string of six histidine residues) by means of which it can be fixed to a chromatographic column (containing, in order to bind His_6-tagged proteins, nickel nitrilotriacetic acid). The column is then used to fractionate a cellular extract and, if binding is reproducible in this system, the interacting protein will bind to the column. The bound proteins are eluted and analysed by SDS-PAGE.

In the 'pull-down' assay (Durfee *et al.*, 1993), one protein is typically expressed fused to a glutathione *S*-transferase moiety (GST). This is incubated with either a tissue extract or the untagged potential interacting protein and glutathione-Sepharose. The GST moiety binds to the glutathione, thereby attaching the bound interacting proteins to the sepharose, which is sedimented by centrifugation. The proteins are removed from the glutathione-sepharose and analysed by SDS-PAGE.

These techniques differ in the method of capturing the interaction complex (i.e. through antibodies recognising the proteins directly, antibodies recognising epitope tags or affinity methods) and in whether a complex native or simple *in vitro* mixture is used as the protein source. None of them, however, can give direct evidence that an interaction occurs within living cells. The strongest evidence for this comes from the yeast two-hybrid system itself. The methods being discussed here merely offer secondary, independent evidence that the proteins physically interact. The co-immunoprecitation method used to verify the RPGR-RPGRIP interaction was successful in this and shows that binding is specific and independent of any additional factors or modifications.

CHAPTER FOUR

CHARACTERISATION OF *RPGRIP*

4.1 - INTRODUCTION

The aims of this work were (i) to isolate the full-length human RPGRIP cDNA, (ii) to determine the chromosomal location of the *RPGRIP* gene, (iii) to describe the genomic structure of *RPGRIP* and (iv) to investigate *RPGRIP* expression in a range of tissues. Finally, (v) the deduced RPGRIP amino acid sequence was analysed *in silico* using a panel of peptide analysis programs.

4.2 - ISOLATION OF THE FULL-LENGTH HUMAN *RPGRIP* cDNA

Alignment of the 10 human and 8 mouse *RPGRIP* ESTs shows that they are clustered, predominantly at the 3' end of the bovine sequence. None of these ESTs provided any additional 5' sequence. The positions of the human ESTs relative to the bovine clone and to each other are shown in Figure 4.1.

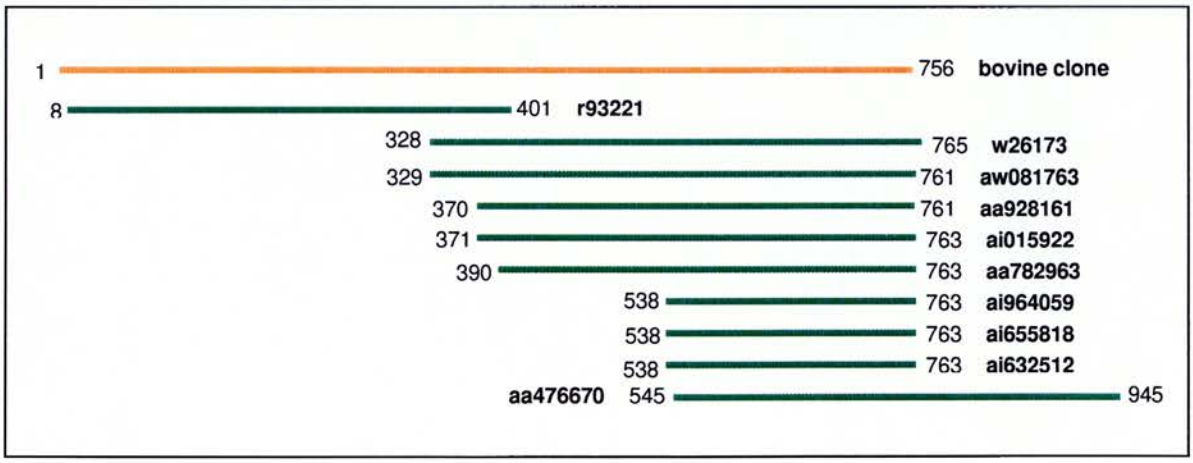


Figure 4.1: Extent of the human *RPGRIP* ESTs (green) relative to the bovine clone (orange) and to each other. The numbers indicate the size and relative extents of the sequences, from 1, at the start of the most 5' sequence (the bovine clone), to 945, at the end of the most 3' sequence (EST aa476670).

It was important to discover if the 5' and 3' ends of the human cDNA represented the true limits of the *RPGRIP* transcript. The second triplet in the bovine clone is an initiation codon and this is preserved in the human open reading frame, suggesting that it may be the start of

the *RPGRIP* transcript. However, sequencing the whole of the cDNA insert of clone r93221 revealed that the open reading frame continues in a 5' direction beyond the first ATG triplet for 28 bp and that there are no stop codons between the vector polylinker and the first ATG. Additional 5' sequence was therefore sought using the 5' rapid amplification of cDNA ends (RACE) technique (see below).

At the 3' end of the aligned sequences, the stop codons of the human, bovine and mouse open reading frames coincide:

```
human   ... ATTTGTTTTC ATGAAGGAAC AAGTGCTATT ...
bovine  ... ATTTGTGTTC CTGATGGAAC AAGTGCTAGT ...
mouse   ... ATCTGTTTTC ATGAAAGTAC AAGTTCTACT ...
```

This suggests that they have been evolutionarily conserved and represent the 3' termini of the open reading frames. 3' RACE experiments were carried out to investigate whether additional 3' exons could be detected.

4.2.1 - 5' Primer extension reactions

Initial attempts to obtain extra upstream cDNA sequence by 5' RACE, using primers complementary to the coding strand, were unsuccessful. This technique was performed using a variety of human RNA samples for first strand cDNA synthesis (including retina poly $(A)^+$ RNA), and a number of different PCR primers and protocols. Commercial adaptor-ligated cDNA libraries were also used for human retina, testis and foetal liver, again without success.

To examine the sequence around the 5' end of the human r93221 clone at the genomic level, another primer extension technique was employed. This was analogous to the RACE method, differing only in that the PCR template DNA is genomic rather than cDNA (Siebert *et al.*, 1995). Five pools of genomic DNA were obtained (Clontech, Basingstoke) that had each been separately digested with restriction endonucleases (*Eco*RV, *Sca*I, *Dra*I, *Pvu*II, and *Ssp*I) and the resulting linear fragments ligated to adaptors at each end. From these samples, it was possible to selectively amplify lengths of genomic DNA using a gene-specific primer and a primer specific for the adaptor DNA. Non-specific amplification from the adaptor primers was prevented by the adaptors having retracted 3' ends that terminate in an amine group. The retracted 3' end eliminated the adaptor primer target (until the extension from the specific primer had extended to the end of the other strand of the adaptor) and the amine group at the free 3' end of the adaptor end blocked 3' extension of this strand and creation of an adaptor primer target site (Figure 4.2). The use of

five distinct genomic DNA samples (each digested with a different endonuclease) increased the likelihood of amplifying a fragment containing the 5' end of *RPGRIP*. Genomic fragments containing the sequence at the 5' end of EST r93221 were amplified in this way. For the first round of PCR, the AP1 adaptor-specific oligonucleotide supplied with the kit was used with 19REV, which is specific for RPGRIP. The second round of amplification was achieved using nested primers: AP2 (adaptor-specific) and 4REV (specific for *RPGRIP*). Specific products were obtained from two of the five templates. These were 'TA' cloned (see section 2.11.4, page 57) and sequenced using M13F and M13R primers. The sequence obtained from these clones extended the open reading frame in a 5' direction by 146 bp (i.e. 146 bp separated the 5' end of the EST r93221 insert from the nearest upstream stop codon).

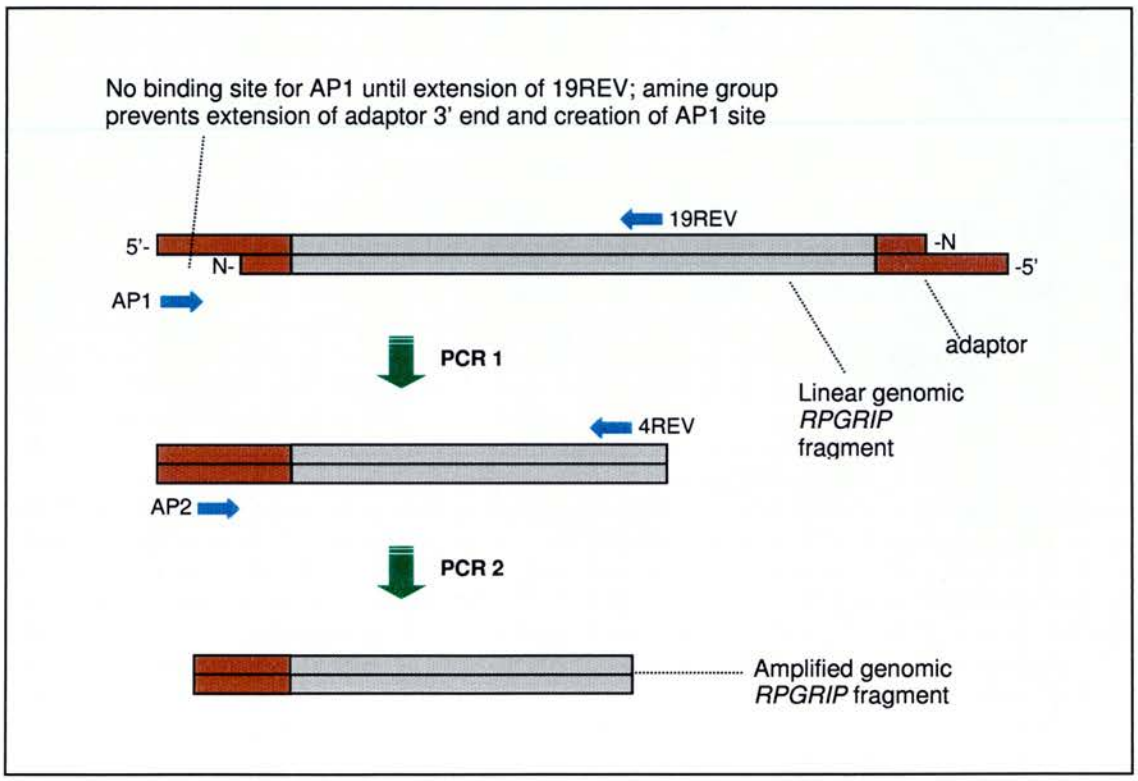


Figure 4.2: Schematic representations of the primer extension technique whereby linear fragments of RPGRIP genomic DNA were amplified by PCR. Grey bars indicate genomic DNA, red bars indicate adaptor sequences. The first round of amplification was primed with the adaptor-specific AP1 and RPGRIP-specific 19 REV oligonucleotides and the second round was primed with the nested adaptor-specific AP2 and RPGRIP-specific 4REV primers.

4.2.2 – 5' RACE

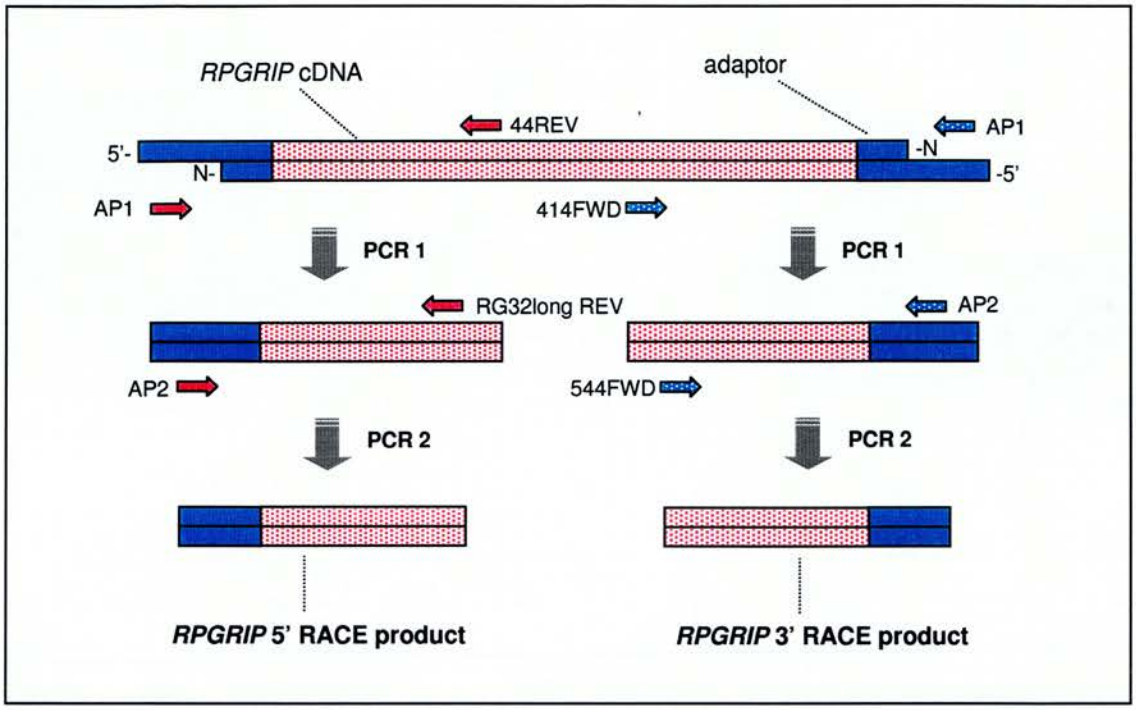


Figure 4.3: Schematic representations of the 5' and 3' RACE technique. Red bars indicate RPGRIP cDNA, blue bars indicate adaptor sequences. The first rounds of PCR-amplification were primed with the adapter-specific AP1 and RPGRIP-specific 44 REV (5') and AP1 and RPGRIP-specific 414 FWD (3') oligonucleotides. The second rounds were primed with the nested adapter-specific AP2 and RPGRIP-specific RG32 long REV (5') and AP2 and RPGRIP-specific 544 FWD.

The new 5' sequence was used to design an oligonucleotide PCR primer, RG32longREV, which was used for a 5' RACE experiment (shown schematically in Figure 4.3) using human retina and testis cDNA pools (Clontech, Basingstoke). The first round of amplification was primed with the AP1 adaptor primer and the *RPGRIP*-specific 44REV primer. The second round of amplification was carried out using nested primers: the AP2 adaptor primer and RG32longREV. The result of this experiment is shown in Figure 4.4.

The specific (though poorly-defined) products in lanes 1 and 4 (indicated by arrows in Figure 4.4) were excised from the gel and purified. The testis product was re-amplified by PCR and sequenced; attempts to re-amplify the retina product were unsuccessful, possibly because of the difficulty in excising a clean band from the agarose gel. The extra 5' sequence obtained from the testis cDNA extended the open reading frame to 1758 bp. The nucleotide context of the first ATG triplet conforms well to the Kozak consensus (Kozak, 1987) as shown in Table 4.1.

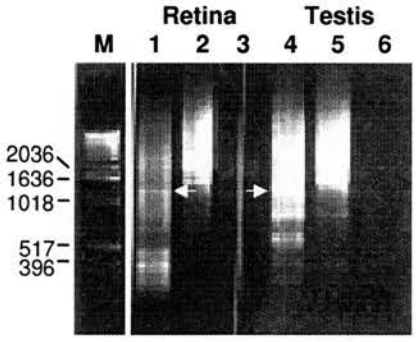


Figure 4.4: Photograph of an agarose gel containing the products of the second round of 5' RACE from retina and testis cDNA (Lane M = marker (sizes in bp), lanes 1 and 4 = products of the PCR primed with both the adaptor- and gene-specific primers, lanes 2 and 5 = adaptor primer only, lanes 3 and 6 = gene specific primer only. The white arrows show the bands that were excised.

Table 4.1: The *hRPGRIP* presumed 5' noncoding sequence from position -9 to position +4 compared with the consensus determined by Kozak (1987). The numbers indicate the percentage of transcripts that have the residue shown in that position.

Position:	-9	-8	-7	-6	-5	-4	-3	-2	-1	+1	+2	+3	+4
Consensus:	G^{33}	C^{39}	C^{37}	G^{44}	C^{39}	C^{53}	A^{61}/G^{36}	C^{49}	C^{55}	A	T	G	G^{46}
***hRPGRIP*:**	A^{23}	T^{19}	T^{20}	G	C	C	A	G^{13}	C	A	T	G	C^{16}

Oligonucleotide PCR primers were designed to anneal to the 5' RACE product (760 delREV and 494 REV, Table 2.1, page 59) and these were used for additional 5' RACE experiments using human retina cDNA (Clontech, Basingstoke). A transcript was identified which starts 12 bp downstream from the start of the testis transcript but no additional 5' coding sequence was obtained. Like the testis 5' RACE product, the retina transcript contains in-frame stop codons upstream of the proposed translation initiation codon, consistent with it containing the first exon.

A BLAST search of the human EST database for sequences similar to the 5' RACE product identified a single EST (accession number w28191) with strong sequence homology (94% identical; e value (the number of different alignments with equal or better scores expected to occur in a search by chance) = 2×10^{-17}). The EST contains *RPGRIP* sequence from nucleotides 408 to 474 and comes from a random primed human retina cDNA library.

4.2.2 - 3' RACE

The 3' end of *hRPGRIP* was amplified from a testis oligo (dT)-primed cDNA pool (Clontech, Basingstoke). The *RPGRIP*-specific oligonucleotides 414FWD and adaptor primer AP1 were used for the first round of amplification and the nested primers 544FWD and AP2 for the second round (see Figure 4.3). A transcript was identified extending 84 bp beyond the end of the open reading frame. The product was sequenced and found to contain

the same in-frame stop codon and no additional 3' coding sequence. 3' RACE experiments using retina cDNA (Clontech, Basingstoke) also failed to detect any additional 3' transcribed sequence. Neither of the two most commonly used hexameric polyadenylation signals (AATAAA and ATTAAA) is present at the 3' end of the testis RACE product or the ESTs. However, 17 bp and 31 bp upstream from the 3' end of the RACE product are two hexamers that differ from the canonical AATAAA by a single nucleotide: TATAAA and AGTAAA respectively. These, according to Graber *et al.* (1999), are the fifth (TATAAA) and fourth (AGTAAA) most commonly used hexameric polyadenylation signals. Of the two, TATAAA but not AGTAAA is conserved in human, bovine and mouse *RPGRIP*. An alignment of these three sequences in the region of the termination codon and the potential polyadenylation signals is shown in Figure 4.5.

```
human_RPGRIP    TGACTGAAGATTTGTTTTCATGAAGGAACAAGTGCTATTCCAATCTAAAAGTCTCTGAGG
bovine_RPGRIP   TGACTGAAGATTTGTGTTCCTGATGGAACAAGTGCTAGTCCATTCTAAA----CCTGAGG
mouse_RPGRIP    TGACTGAAGATCTGTTTTCATGAAAGTACAAGTTCTACTCCACTCAACGA---TTTGAGG

human_RPGRIP    GAACCATAGTAAAAAGTCTCTTATAAAGTTAGCTTGCTATAACA
bovine_RPGRIP   GAACCATAGTCAAAAGTCTCTTATAAAGT-AACTTGCTCTACTA
mouse_RPGRIP    GCATCACAGTCAAA-GTGTCTTATAAAAT-AACATGACATGCCA
```

Figure 4.5: CLUSTAL W alignment (Thompson *et al.*, 1994) of human, bovine and mouse *RPGRIP* 3' sequences. The conserved termination codon is boxed (green), as is a conserved potential polyadenylation signal (red) and a potential polyadenylation signal present in human *RPGRIP* only (pink). Blue letters indicate conserved nucleotides in these positions, black letters represent non-conserved nucleotides.

4.3 - GENOMIC STRUCTURE OF *hRPGRIP*

Prior to the release of 196 kb of chromosome 14 genomic sequence containing *hRPGRIP* (see below), much of the genomic structure was determined by sequencing PCR products generated using genomic DNA and different combinations of RPGRIP cDNA-specific primers. Fourteen intron/exon boundaries were established in this way using the following combinations of PCR primers (the sequences of which are given in Table 2.1, page 59): 430 FWD and 494 REV to determine the exon 2/intron 2 and intron 2/exon 3 boundaries, 801 FWD and Long REV to determine the exon 8/intron 8, intron 8/exon 9, exon 9/intron 9 and intron 9/ exon 10 boundaries, 139 FWD and 183 REV to determine the exon 11/intron 11 and intron 11/exon 12 boundaries, 419 FWD and 522 REV to determine the exon 12/intron 12 and intron 12/exon 13 boundaries, 414 FWD and 544 REV to determine the exon

13/intron 13 and intron 13/exon 14 boundaries, 521 FWD and 761 REV to determine the exon 14/intron 14 and intron 14/exon 15 boundaries (illustrated in Figure 4.6).

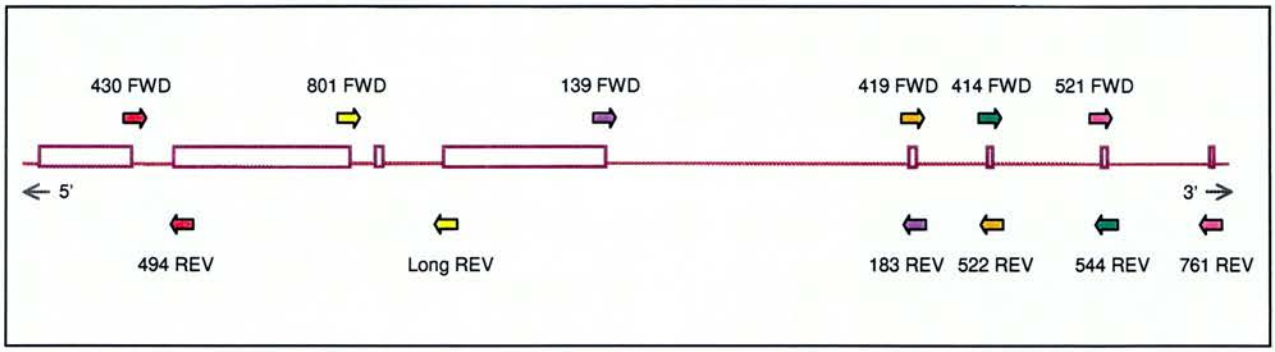


Figure 4.6: Generation of PCR products used to determine fourteen *RPGRIP* intron/exon boundaries (open boxes = exonic *RPGRIP*, line = intronic *RPGRIP*, arrows of same colour show forward and reverse PCR primers used to amplify segments of the *RPGRIP* gene).

The remaining boundaries were obtained, towards the end of this project, following the database appearance of 196 kb of genomic sequence from a chromosome 14 contig (accession number AL135744). The open reading frame that was assembled from the 5' RACE product and the r93221 clone is divided into 12 exons (although see below). The boundaries conform to the GT/AG rule (Mount, 1982) with the exception of the splice donor site at the end of exon 4. The sequence at the exon 4/intron 4 junction was repeatedly checked and confirmed as C T G **g c** a a g t (the exonic sequence is in upper case letters; the intronic sequence is in lower case letters). It is rare to observe a GC instead of the conserved GT present at most splice donor sites but other examples of functional splice sites with this sequence have been reported (Senapathy, Shapiro and Harris, 1990). The exon/intron boundaries and intron sizes are shown in Table 4.2 (page 124).

The HGMP RC NIX panel of bioinformatics programs (http://www.hgmp.mrc.ac.uk/Registered/Webapp/nix/) was applied to the 196 kb genomic sequence containing *RPGRIP* (AL135744). The GENSCAN exon-recognition program (Burge and Karlin, 1998) predicted the presence of three previously undetected exons in the middle of intron 4 and twelve exons upstream of exon 1 that were absent from 5' RACE products. The splice sites and exon and intron sizes of the twelve upstream GENSCAN-predicted exons are given in Table 4.3 (page 125) and their sequences are given in Table 4.4 (page 126). The relative positions of these GENSCAN-predicted exons (numbered -12 to -1 and 5 to 7; blue) and the previously-identified *RPGRIP* exons (numbered 1 to 4 and 8 to 15; red) are shown in Figure 4.7.

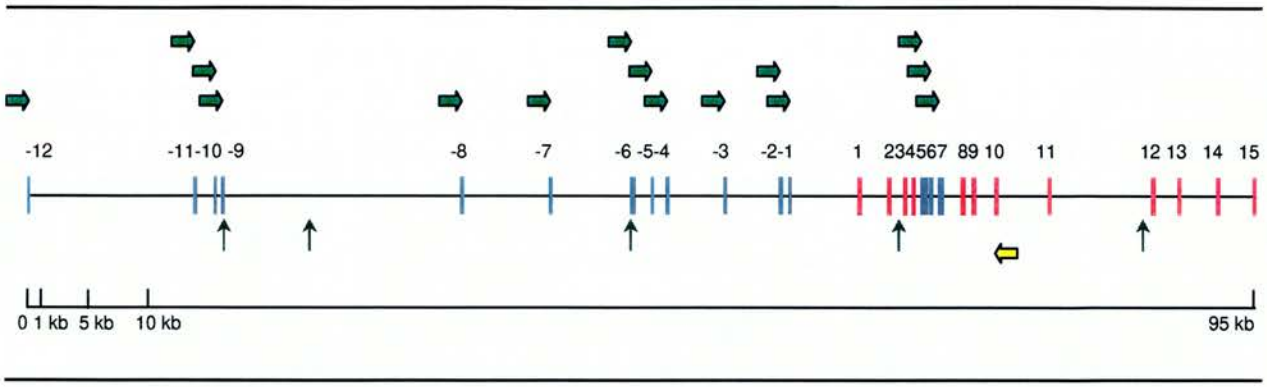


Figure 4.7: Exons predicted by GENSCAN analysis of the chromosome 14 genomic sequence (AL135744) containing *RPGRIP*. Previously-identified *RPGRIP* exons are shown in red and additional exons are shown in blue. Horizontal green arrows indicate the positions of GENSCAN-predicted exon-specific PCR primers; the horizontal yellow arrow indicates the target site for the *RPGRIP* exon 10-specific PCR primer RG32longREV. Vertical arrows indicate the positions of GrailEXP-predicted CpG islands.

As a component of the NIX analysis, the chromosome 14 sequence (AL135744) was examined using the GrailEXP program (version 2.0; Hyatt *et al.*, 2000) in an attempt to identify CpG islands (see page 151). Six potential CpG islands were identified (see Figure 4.7): (1) 699 bp of DNA located 111 bp downstream of GENSCAN-predicted exon -10 and encompassing the whole of exon -9, (2) 357 bp of DNA located 6,638 bp downstream of GENSCAN-predicted exon -9 and 11,447 bp upstream of GENSCAN-predicted exon -8, (3) 259 bp encompassing most of GENSCAN-predicted exon -6 (beginning 12 bp upstream of the start of the predicted exon and ending 27 bp upstream of the end of it), (4) 265 bp located 672 bp upstream of exon 3, (5) 288 bp located 1,613 bp upstream of exon 12 and (6) 1061 bp located 31,605 bp downstream of the terminal exon (exon 15). TATA, GC and CAAT boxes were not found in the immediate vicinity of *RPGRIP* exon 1, or any of the GENSCAN-predicted exons.

Attempts were made to PCR-amplify the GENSCAN-predicted exons from human retina-derived cDNA using primers designed to anneal to these exons (Genscan exon 1-12 and 17-19,) and a primer specific for a previously identified *RPGRIP* exon (RG32longREV). The annealing site locations of the GENSCAN-predicted exon-specific primers (green) and RG32longREV (yellow) are shown schematically in Figure 4.7 and their sequences are given in Table 2.1 (page 59). The 3 exons predicted to lie in intron 4 (GENSCAN-predicted exons 5 to 7) were detected and maintained the open reading frame (ORF) of the *RPGRIP* transcript, but the 12 upstream exons (GENSCAN-predicted exons -12 to -1, which are also in a continuous ORF with the rest of *RPGRIP*) were not detected (Figure 4.8). This

indicated that *RPGRIP* consists of 15 exons: the twelve previously identified exons (exons 1-4 and 8-15) and the three GENSCAN-predicted exons downstream of exon 4 (exons 5-7).

Table 4.2: Genomic structure of the human *RPGRIP* gene. Exonic sequence is in upper case letter, genomic sequence in lower case. The atypical splice donor site at the exon-intron junction at the end of exon 4 is underlined.

Exon		Sequence at intron-exon junction		Intron	
Number	Size	Splice acceptor	Splice donor	Number	Size
5' UTR	104	tatcacacccttggggtatccatctgagagcttgc tttccattgccagcATGCTGGACA			
1	154		AGGGAGAAAGgtaggctggaccttgaag agctctctcaaatgtggcatcctctacctctg	1	2165
2	161	tggtgataaataactacagaatttcacatttctggatt atttttccccagCCCAAAATGA	TCAGATCGAGgtaagagcctctttaaaca aactagtccactctgcagagagatctctggc	2	1080
3	144	tttacctgtcatatttatactccctttaccaatgcgttt ccctctacagCCAAGTGAAC	GTGTTATCAGgtgcaaggaaagatggta caggaaggggatggataacaggaacgtgggaa	3	1080
4	151	ttggttttaggccactgagatagaaaagttcagaca ttattttgtttcagGAGGAACTGG	CCAACATCTGgcaagtcttagtcctttgttc tcctcacttcgggacccttccacagctaa	4	2612
5	453	gaaagagctccctacccttttaacggataggcagct ttctttcccctctagAACAGCTCAAA	ACACTGATTGgtaagtgccgttggcttcct gcggctcctaagcaccaatgcagaatttcc	5	161
6	152	cttgccacaccatctgtattccccctgctttcacactt tctgctacccagGAGCTGGTGG	GGAAGAGGAGgtgagaaaaaagatgtg ccgaggcatctcagaggagcctcagccaaacag	6	445
7	343	aacagtctcaagctgccctttttcctcaatccatgac caacatctttccagTTCAGATCGG	TCTATCAAAGgtgggagttcgaggttatta catcttcacgccctcttcccagtgttgtct	7	1437
8	185	cttgcttatttcatgtgatcaggtcttattaatatctgtt tgtttctcagGTGATTTTAA	TCCTTCCCAGgtaactctccaggactcca caggtagcagatctctgccaatcctatggag	8	612
9	204	aggttgttaaactaccagcttgtaatgctatctctgat ctttctattcagGATCAGATGG	TACACCAGAGgtaagaccttaaaaactct gaagcactaaattgtgggttgtagtgtcccc	9	1618
10	139	gaaaaccaagatattaccagctatgtagtattgttgt tcttattctgaagCAGGTGAATT	CCTGTAAATGgtattgtcttttaaaatctattt tttttcctagcactttggggaggccga	10	4215
11	101	ggctaaagtgctttgaaacagttctataactgcaac ctcttctctagcagACAAAGAATC	TCCTAAGGCAgtaagtacactggagtaat cattgcatacgagataagacgataaataatg	11	8329
12	193	cttggagcctcactaacctttaggaactaaataaac attttccttatcagGATTCAGAGA	TTTAGCAAGGgtgaggcatcctgtgtggt tactggggtgaggaagtctgatgaacattga	12	1884
13	85	cactgcaacagtatatgattctttgctttttctcttac ccttaatacagTAATAGACCT	ATCAAGGACAgtaagcatctgctttccact ttgaaacaaaggagatattgagacagaaat	13	2973
14	131	gaagcattaagagtatcaacagtgctgaattaaatg caatttctttttagTTTAAAGTTT	GAGCTAGACAgtgagtcatttttttttcagtt ctaattatttccaaggaaatcatcctaa	14	2800
15	113	gtatcaaagtgttcaactgagtgatgctgtttttttcc ctttcccaacagTTGTTAGCCC	TTTGTTTTCATGAaagtgctattccaatc taaaagtctctgagggaaccatagtaaaaggaac	15	

Table 4.3: Genomic structure of the hypothetical portion of *hRPGRIP* predicted using the GENSCAN program (Burge and Karlin, 1998) to lie upstream of exon 1. Attempts to amplify these exons from human retina cDNA were unsuccessful. Exonic sequence is in upper case letter, genomic sequence in lower case.

Exon		Sequence at intron-exon junction		Intron	
Number	Size	Splice acceptor	Splice donor	Number	Size
5' UTR		tggaagaacaccccaaatgttatctttggatagctg tattttctaaaaATGTTCTATG			
-12	38		CGGATCAAAGgtaagataaagatgaggt acctcctctcccaaatccacatcaatcttcag	-12	12888
-11	145	attagagcaactagctattttggagtttattctttgctt gtttcgcacagCCATTCCATC	GAAAAAAAAGgtcagaaaacctgcaac ctggccatccttcttcaacttttccccctagcc	-11	1359
-10	186	atctacttggtaaccttcccccttcccctccccact ttttacatgtaagAAGAACGGAG	AGGACACCAGgttaaacctatacagagg ctgaggcctataaaactgcccccggcttcaac	-10	245
-9	142	gagggagaagagattcgattctgagtctcctactc ccgggttctgcgtagAGAAGCCGAC	CCTATTGGCTgtgcgtttttcgctcccgct cctaattagacagttcaaccgctcctcac	-9	18441
-8	123	aagacatcctaaagttgcatgtaactgactggactt ttcctttatttcagTGTCCTCTGG	GCCTCAAAAGgtaacttctacgcctgagg atggacacctttagaattaaaagaactctc	-8	6616
-7	133	aaaaaataggtgtaaaaatactgtccatttactgtttt ctgggactttagGTAAGAATAT	AGATCAAAAGgtacttagagttctccttaa attttttttttttttttttttttttttttt	-7	6156
-6	272	aaataataactgtcatgaaaggagaagccacgtg ctctatttgtccacagGCTGAGGACC	CCCAAGAGGGgtgagatttaaggctacat cccctacagggctaagacactgggaaggact	-6	1249
-5	97	ttccggagggtactgttgcccgagtctgcatcttct gtctaaacttttagGGCCAAGGGA	CCAGTGAACTgtgagttcttctcatttatcc catgttgttattcctgccactttaattag	-5	828
-4	132	agcagtgtcataagtatggctaaacccattggtcta tgcatgcctaacagTGCCCACATC	CAGAGCTTCGgtaagagtgtggcactcca tgcctcaaaccaaaacgaactgaaaaagcta	-4	4188
-3	106	tttaattctatccatgttcagacagaataatttagcgc ctttctctgcagAGCTTCCATT	AGTTCAAGAGgtgagttgccatcatcagc tgtgctttcttggtgggggggaaaccccaatt	-3	3986
-2	147	tcgtgctgagtgatatgaccattcccagaggtacttt ccttttgacccagAATCAGGGAA	TCTGAAGGAGgtaaataataatagttgaa agataccatctacatttcaacttcagggatg	-2	469
-1	68	ggggaatagattttaacattttatctcaagggctact atcactcttagtttcagGAGAGAGTTG	TCTTAGAAAGgtgagtaccacatttgggt cccagagcagtgttactatgaaagacattat	-1	5191

Table 4.4: The 12 exons predicted using the GENSCAN program (Burge and Karlin, 1998) to lie upstream of the first detected *RPGRIP* exon. Attempts to amplify these exons from human retina cDNA were unsuccessful.

Exon	Sequence
-12	ATGTTCTATGAGCATGTATTACTTTATTCGGATCAAAG
-11	CCATTCCATCCAAAACAAGGTCGAATGCCATGTCAAATCCTCTAGAGATTTAATTACAAACGTAAGCTG TTTAAATTCTCTACAGAAGGAACATCAACCACCAAGTATTCCCGAGAGTGGCACTAGCATGAGGGGGA AAAAAAAG
-10	AAGAACGGAGGTAATAGAAGGCATCGGCGCTACGTTAACTTCGACTATATTAATGAGCATCACGATCTT CATTTTCATGATCCCGTAGTCCTCGCGGTCGTAGAAAATGAAAGGAAAGCAGCCGTTAACACAGGAAG AGCGAAAAAAGGACATTTTCCCTGGCGATCGTGGCCTAGAGGACACCAG
-9	AGAAGCCGACTGCTGCTGGAGGTCGGCAACGCGGCCACAACCGCTCAGTCTTCGTCTCTTCACAAAATG GCCCCCGAGTCTCCTCTGCCGCCGAGGCGAAGGCCTAGTCCCCGCCCCATTCCTGAGAATCTTCCTATTG GCT
-8	TGTCCTCTGGGATCTCTTACAGCTTGGGAACAGAGATCATGTCACATCTGGTGGACCCTACATCAGGAG ACTTGCCAGTTAGAGACATAGATGCTATACCTCTGGTGCTACCAGCCTCAAAAG
-7	GTAAGAATATGAAAACTCAACCACCCTTGAGCAGGATGAACCGGGAGGAATTGGAGGACAGTTTCTTT CGACTTCGCGAAGATCACATGTTGGTGAAGGAGCTTTCTTGGAAGCAACAGGATGAGATCAAAAG
-6	GCTGAGGACCACCTTGCTGCGGTTGACCGCTGCTGGCCGGGACCTGCGGGTCGCGGAGGAGGCGGCGC CGCTCTCGGAGACCGCAAGGCGCGGGCAGAAGGCGGGATGGCGGCAGCGCCTCTCCATGCACCAGCGC CCCCAGATGCACCGACTGCAAGGGCATTTCCACTGCGTCGGCCCTGCCAGCCCCCGCCGCGCCCAGCCT CGCGTCCAAGTGGGACACAGACAGCTCCACACAGCCGGTGCACCGGTGCCGGAGAAACCCAAGAGGG
-5	GGCCAAGGGACAGGCTGAGCTACACAGCCCCTCCATCGTTTAAGGAGCATGCGACAAATGAAAACAGA GGTGAAGTAGCCAGTAAACCCAGTGAACT
-4	TGCCCACATCATGGCCAGCAATACCATGCAAGTGGAAGAGCCACCCAAGTCTCCTGAGAAAATGTGGC CTAAAGATGAAAATTTTGAACAGAGAAGCTCATTGGAGTGTGCTCAGAAGGCTGCAGAGCTTCG
-3	AGCTTCCATTAAAGAGAAGGTAGAGCTGATTCGACTTAAGAAGCTCTTACATGAAAGAAATGCTTCATT GGTTATGACAAAAGCACAATTAACAGAAGTTCAAGAG
-2	AATCAGGGAATCCTGAGTGCAGCCCATGAGGCCCTCCTCAAGCAAGTGAATGAGCTCAGGGCAGAGCT GAAGGAAGAAAGCAAGAAGGCTGTGAGCTTGAAGAGCCAACTGGAAGATGTGTCTATCTTGCAGATGA CTCTGAAGGAG
-1	GAGAGAGTTGAAGATTTGGAAAAAGAACGAAAATTGCTGAATGACAATTATGACAAACTCTTAGAAAG

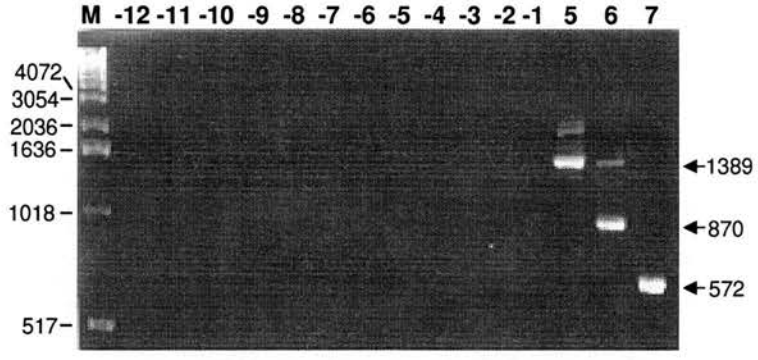


Figure 4.8: Amplification of RPGRIP fragments containing GENSCAN-predicted exons from human retina cDNA. M= marker lane (sizes in bp). Lanes -12 to -1 = products of PCR reactions primed with oligonucleotides specific for GENSCAN-predicted exons -12 to -1 (upstream of *RPGRIP* exon 1) and a primer specific for a previously identified *RPGRIP* exon (RG32longREV); lanes 5 to 7 = products of PCR reactions primed with oligonucleotides specific for GENSCAN-predicted exons 5 to 7 (located downstream of exon 4) and RG32longREV. The 1389 bp product in the '5' lane includes exons 8-10 and GENSCAN-predicted exons 5-7. The 870 bp product in the '6' lane includes exons 8-10 and GENSCAN-predicted exons 6-7. The 572 bp product in the '7' lane includes exons 8-10 and GENSCAN-predicted exon 7.

The precise boundaries of the three additional exons (exons 5-7) were determined by sequencing across them at the cDNA level. To achieve this, the 1389 bp lane '5' band shown in Figure 4.8 was sequenced as was a PCR product generated from human retina cDNA using an exon 4-specific forward primer (638 FWD long, Table 2.1, page 59) and an exon 5-specific reverse primer (?4/5 junction REV, Table 2.1, page 59). The boundaries were found to coincide with the GENSCAN prediction. Examination of these exons at the genomic level (using the AL135744 chromosome 14 sequence submission) revealed these exons to have canonical splice sites (shown in Table 4.3). It is interesting to note that these are the three exons immediately 3' to the non-canonical splice site at the exon 4/intron 4 junction splice site and are absent from PCR-amplified *RPGRIP* transcripts.

Interestingly, PCR-amplification of *RPGRIP* from human testis cDNA using primers in the 5' and 3' untranslated regions (UTR) (-26F and Stop +36, Table 2.1, page 59) produced 2 bands: one of the expected size of 1862 bp and a larger but much less abundant band of approximately 2350 bp (Figure 4.9). This estimated 488 bp of additional sequence is large enough to include either exon 5 (453 bp) or exons 6 and 7 together (495 bp). However, repeated attempts to purify and 'TA' clone DNA from the larger band were unsuccessful.

The extra exons lengthen the cDNA from 1,946 bp to 2,894 bp and the open reading frame from 1,758 bp to 2,706 bp, predicting a protein of 902 amino acids and 97 kD. The

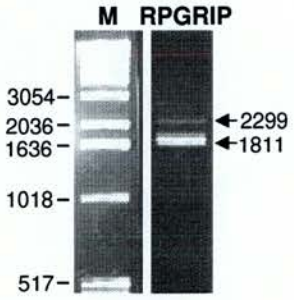


Figure 4.9: RPGRIP amplified by PCR from testis cDNA using primers in the 5' and 3' UTRs produced 2 products of different sizes (this photo shows the slightly (51 bp) smaller products of a second round of amplification using nested primers (-6 Fwd and stop+6). M= marker (sizes in bp).

full-length *hRPGRIP* cDNA with intron/exon boundaries is shown in Figure 4.11. The sequence was submitted to GenBank and assigned the accession number AF260257. The *RPGRIP* gene covers 33,495 bp of genomic DNA and consists of 15 exons. The average exon size is 188 bp (range = 85 – 453). The average intron size is 2199 bp (range = 161 – 8329). A schematic representation of the *hRPGRIP* genomic structure is shown in Figure 4.10.

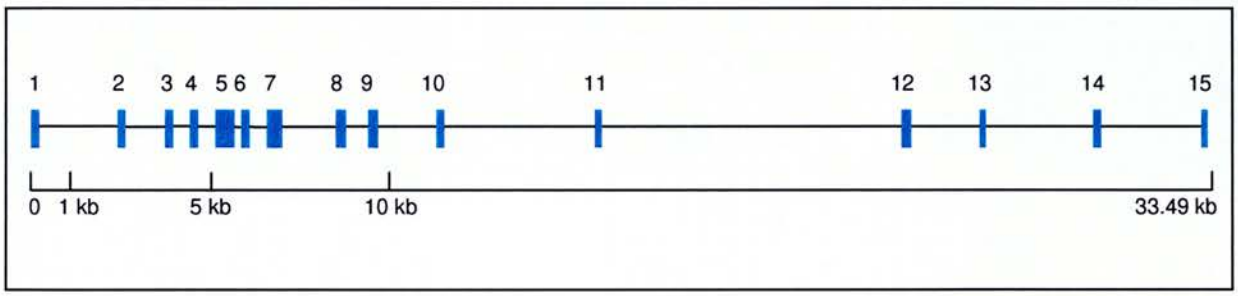
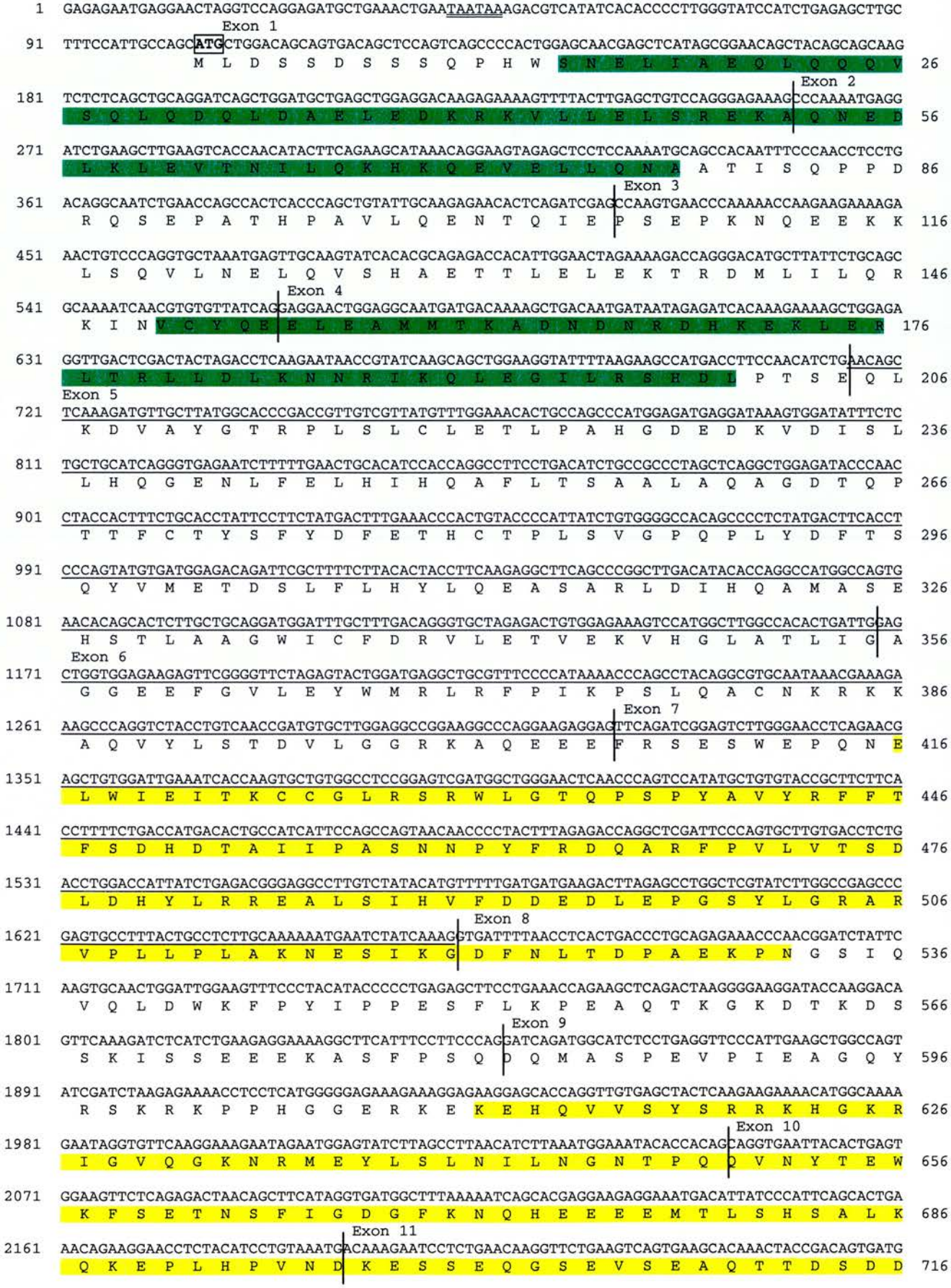


Figure 4.10: Schematic representation of the genomic structure of *hRPGRIP*. The numbered blue boxes indicate the relative positions and sizes of *RPGRIP* exons 1 to 15.


1 GAGAGAATGAGGAACTAGGTCCAGGAGATGCTGAAACTGAATAATAAAGACGTCATATCACACCCCTTGGGTATCCATCTGAGAGCTTGC
Exon 1
91 TTTCCATTGCCAGCATGCTGGACAGCAGTGACAGCTCCAGTCAGCCCCACTGGAGCAACGAGCTCATAGCGGAACAGCTACAGCAGCAAG
M L D S S D S S S Q P H W S N E L I A E Q L Q Q Q V 26
Exon 2
181 TCTCTCAGCTGCAGGATCAGCTGGATGCTGAGCTGGAGGACAAGAGAAAAGTTTTACTTGAGCTGTCCAGGGAGAAAGCCCAAAATGAGG
S Q L Q D Q L D A E L E D K R K V L L E L S R E K A Q N E D 56
271 ATCTGAAGCTTGAAGTCACCAACATACTTCAGAAGCATAAACAGGAAGTAGAGCTCCTCCAAAATGCAGCCACAATTTCCCAACCTCCTG
L K L E V T N I L Q K H K Q E V E L L Q N A A T I S Q P P D 86
Exon 3
361 ACAGGCAATCTGAACCAGCCACTCACCCAGCTGTATTGCAAGAGAACACTCAGATCGAGCCAAGTGAACCCAAAAACCAAGAAGAAAAGA
R Q S E P A T H P A V L Q E N T Q I E P S E P K N Q E E K K 116
451 AACTGTCCCAGGTGCTAAATGAGTTGCAAGTATCACACGCAGAGACCACATTGGAACTAGAAAAGACCAGGGACATGCTTATTCTGCAGC
L S Q V L N E L Q V S H A E T T L E L E K T R D M L I L Q R 146
Exon 4
541 GCAAAATCAACGTGTGTTATCAGGAGGAACTGGAGGCAATGATGACAAAAGCTGACAATGATAATAGAGATCACAAAGAAAAGCTGGAGA
K I N V C Y Q E E L E A M M T K A D N D N R D H K E K L E R 176
631 GGTTGACTCGACTACTAGACCTCAAGAATAACCGTATCAAGCAGCTGGAAGGTATTTTAAGAAGCCATGACCTTCCAACATCTGAACAGC
L T R L L D L K N N R I K Q L E G I L R S H D L P T S E Q L 206
Exon 5
721 TCAAAGATGTTGCTTATGGCACCCGACCGTTGTCGTTATGTTTGGAAACACTGCCAGCCCATGGAGATGAGGATAAAGTGGATATTTCTC
K D V A Y G T R P L S L C L E T L P A H G D E D K V D I S L 236
811 TGCTGCATCAGGGTGAGAATCTTTTTGAACTGCACATCCACCAGGCCTTCCTGACATCTGCCGCCCTAGCTCAGGCTGGAGATACCCAAC
L H Q G E N L F E L H I H Q A F L T S A A L A Q A G D T Q P 266
901 CTACCACTTTCTGCACCTATTCCTTCTATGACTTTGAAACCCACTGTACCCCATTATCTGTGGGGCCACAGCCCCTCTATGACTTCACCT
T T F C T Y S F Y D F E T H C T P L S V G P Q P L Y D F T S 296
991 CCCAGTATGTGATGGAGACAGATTCGCTTTTCTTACACTACCTTCAAGAGGCTTCAGCCCGGCTTGACATACACCAGGCCATGGCCAGTG
Q Y V M E T D S L F L H Y L Q E A S A R L D I H Q A M A S E 326
1081 AACACAGCACTCTTGCTGCAGGATGGATTTGCTTTGACAGGGTGCTAGAGACTGTGGAGAAAGTCCATGGCTTGGCCACACTGATTGGAG
H S T L A A G W I C F D R V L E T V E K V H G L A T L I G A 356
Exon 6
1171 CTGGTGGAGAAGAGTTCGGGGTTCTAGAGTACTGGATGAGGCTGCGTTTCCCCATAAAACCCAGCCTACAGGCGTGCAATAAACGAAAGA
G G E E F G V L E Y W M R L R F P I K P S L Q A C N K R K K 386
Exon 7
1261 AAGCCCAGGTCTACCTGTCAACCGATGTGCTTGGAGGCCGGAAGGCCCAGGAAGAGGAGTTCAGATCGGAGTCTTGGGAACCTCAGAACG
A Q V Y L S T D V L G G R K A Q E E E F R S E S W E P Q N E 416
1351 AGCTGTGGATTGAAATCACCAAGTGCTGTGGCCTCCGGAGTCGATGGCTGGGAACTCAACCCAGTCCATATGCTGTGTACCGCTTCTTCA
L W I E I T K C C G L R S R W L G T Q P S P Y A V Y R F F T 446
1441 CCTTTTCTGACCATGACACTGCCATCATTCCAGCCAGTAACAACCCCTACTTTAGAGACCAGGCTCGATTCCCAGTGCTTGTGACCTCTG
F S D H D T A I I P A S N N P Y F R D Q A R F P V L V T S D 476
1531 ACCTGGACCATTATCTGAGACGGGAGGCCTTGTCTATACATGTTTTTGATGATGAAGACTTAGAGCCTGGCTCGTATCTTGGCCGAGCCC
L D H Y L R R E A L S I H V F D D E D L E P G S Y L G R A R 506
Exon 8
1621 GAGTGCCTTTACTGCCTCTTGCAAAAAATGAATCTATCAAAGGTGATTTTAACCTCACTGACCCTGCAGAGAAACCCAACGGATCTATTC
V P L L P L A K N E S I K G D F N L T D P A E K P N G S I Q 536
1711 AAGTGCAACTGGATTGGAAGTTTCCCTACATACCCCCTGAGAGCTTCCTGAAACCAGAAGCTCAGACTAAGGGGAAGGATACCAAGGACA
V Q L D W K F P Y I P P E S F L K P E A Q T K G K D T K D S 566
Exon 9
1801 GTTCAAAGATCTCATCTGAAGAGGAAAAGGCTTCATTTCCTTCCCAGGATCAGATGGCATCTCCTGAGGTTCCCATTGAAGCTGGCCAGT
S K I S S E E E K A S F P S Q D Q M A S P E V P I E A G Q Y 596
1891 ATCGATCTAAGAGAAAACCTCCTCATGGGGGAGAAAGAAAGGAGAAGGAGCACCAGGTTGTGAGCTACTCAAGAAGAAAACATGGCAAAA
R S K R K P P H G G E R K E K E H Q V V S Y S R R K H G K R 626
Exon 10
1981 GAATAGGTGTTCAAGGAAAGAATAGAATGGAGTATCTTAGCCTTAACATCTTAAATGGAAATACACCACAGCAGGTGAATTACACTGAGT
I G V Q G K N R M E Y L S L N I L N G N T P Q Q V N Y T E W 656
2071 GGAAGTTCTCAGAGACTAACAGCTTCATAGGTGATGGCTTTAAAAATCAGCACGAGGAAGAGGAAATGACATTATCCCATTCAGCACTGA
K F S E T N S F I G D G F K N Q H E E E E M T L S H S A L K 686
Exon 11
2161 AACAGAAGGAACCTCTACATCCTGTAAATGACAAAGAATCCTCTGAACAAGGTTCTGAAGTCAGTGAAGCACAAACTACCGACAGTGATG
Q K E P L H P V N D K E S S E Q G S E V S E A Q T T D S D D 716

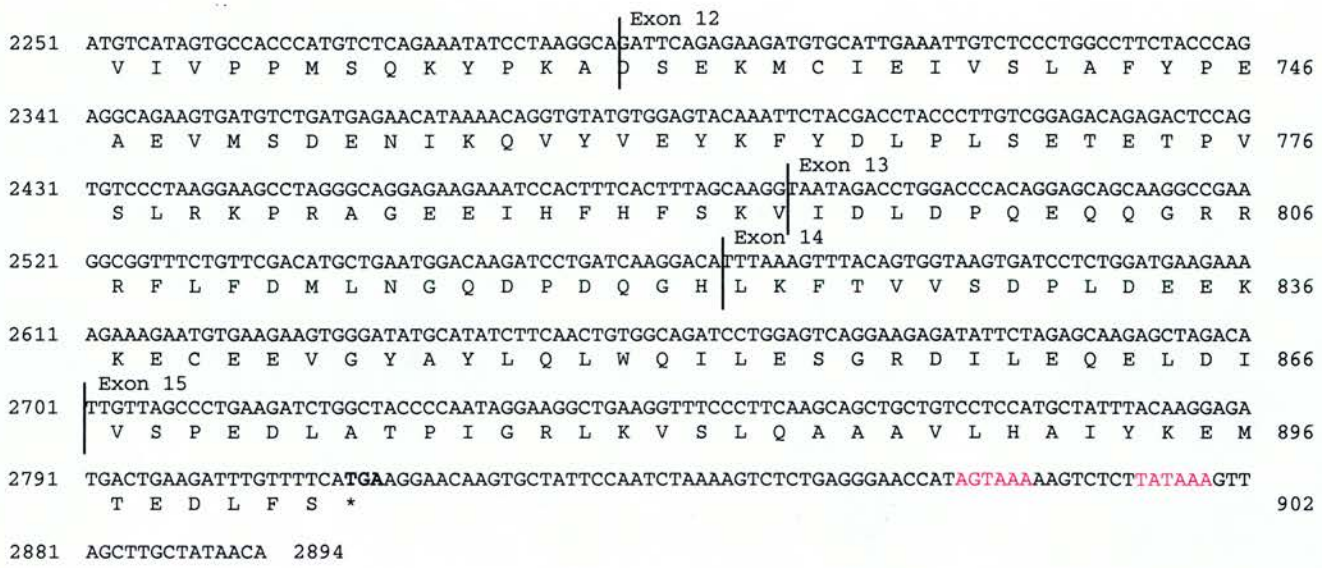


Figure 4.11: cDNA and predicted amino acid sequence of the *RPGRIP* gene (Acc. no. AF260257). The translation initiation codon is boxed and in-frame upstream stop codons are double underlined. The stop codon at the 3' end of the sequence is in bold and two potential polyadenylation signals are shown in red. Predicted coiled-coil (green) and C2 domains (yellow) are highlighted (see section 4.6.1) and the three alternatively spliced exons (exons 5-7) are underlined. Exon-intron boundaries are shown.

4.4 - CHROMOSOMAL MAPPING OF *hRPGRIP*

Four separate lines of evidence were used to determine the chromosomal location of the *hRPGRIP* gene. These were (i) screening a monochromosomal somatic cell hybrid DNA panel, (ii) radiation hybrid mapping, (iii) the identification of an STS containing the *RPGRIP* sequence and (iv) the identification of a GenBank sequence submission containing *hRPGRIP*.

4.4.1 - Screening a monochromosomal somatic cell hybrid DNA panel

A panel of human monochromosomal somatic cell hybrid DNAs was obtained from the MRC HGMP Resource Centre (Kelsell, *et al.*, 1995). This was screened by PCR using the *RPGRIP*-specific 419FWD and 522REV primers to determine the chromosomal location of *RPGRIP*. The human chromosome 14-containing hybrid sample was found to provide a template for amplification of *RPGRIP* while none of the other panel samples, including the rodent control, produced fragments of the correct size. This result is shown in Figure 4.12.

The chromosome 14 sample consists of chromosome 14 DNA plus part of chromosome 16, most likely to be 16p13.1-16q22.1 (HGMP RC, product information leaflet). The negative result obtained with chromosome 16 DNA indicated that *RPGRIP* is located on chromosome 14.

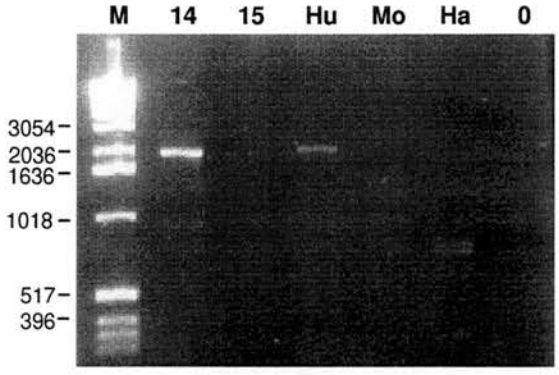


Figure 4.12: Screening a mono-chromosomal somatic cell hybrid DNA panel. (M=Marker lane (sizes in bp), 14=human chromosome 14 DNA, 15=human chromosome 15 DNA, Hu=total human genomic DNA, Mo=total mouse genomic DNA, Ha=total hamster genomic DNA, 0=no DNA.) Hybrids containing all other human chromosomes were negative.

4.4.2 - Identification of an STS containing *RPGRIP* sequence

A BLAST search of the STS database for sequences with homology to *RPGRIP* identified STS:h14a1407 (accession number G35988). This contains 320 bp of *RPGRIP* genomic sequence (covering the intron 12 - exon 13 boundary) which had been localised to chromosomal region 14q11.1 by Dear *et al.* (1998).

The location of h14a1407 was determined during the generation of a high density (100 kb resolution) STS map of the chromosome 14. The method used to do this is called "HAPPY mapping" and is an *in vitro* technique similar in principle to radiation hybrid mapping that replaces radiation hybrids with samples of pure genomic DNA (Dear *et al.*, 1998). Genomic DNA is broken into a pool of fragments by shearing or irradiation and 96 samples are taken from it, each containing less than one whole genome: this constitutes the mapping panel. The more closely linked two markers are, the more likely they are to localise to the same DNA fragment. The mapping panel is screened by PCR in order to determine which STS markers are present in each of the 96 samples. The results of this are used to generate a map, with distances between markers estimated from the frequency with which they co-segregate.

4.4.3 - Radiation Hybrid mapping

In order to determine a more precise chromosomal location for *RPGRIP*, a PCR analysis was performed on the HGMP-RC subset of the Genebridge 4 radiation hybrid panel (Gyapay *et al.,* 1996). The PCR primers used were the same as for screening the monochromosomal somatic cell hybrid DNA panel. The results were submitted to the Whitehead Institute and indicated that *RPGRIP* is located between markers D14S275 and D14S264 with a LOD score of 2.03 relative to the next most likely placement (6.94 centiRays from D14S264) in 14q11.

4.4.4 – GenBank submission

The location of RPGRIP was independently confirmed as being on chromosome 14 when a 196 kb sequence submission (accession number AL135744) from a bacterial artificial chromosome (BAC) clone containing human chromosome 14 DNA including the entire *hRPGRIP* gene sequence appeared in the GenBank database.

4.5 - EXPRESSION OF *RPGRIP*

To assess the expression of *RPGRIP* in various tissues, northern blots were probed, reverse transcriptase-PCR (RT-PCR) experiments were carried out, and *in situ* hybridisation of labelled RPGRIP cDNA to embryonic mouse sections was performed.

4.5.1 - Northern blots

Three northern blots were probed using radiolabelled RPGRIP cDNA to determine the size and tissue distribution of RPGRIP transcripts.

4.5.1.i - Northern blot 1

A commercial human multiple tissue northern blot (MTN2, Clontech, Basingstoke) was probed with a 713 bp ^{32}P-labelled *RPGRIP* fragment corresponding to RPGRIP nucleotides 2135 to 2847 (from the middle of exon 10 to the end of exon 15, see section 2.22, page 69). Two different transcripts were detected in the testis lane: a strong band of 2.0 kb and a weaker band of 3.1 kb, but no signal was detected in lanes containing spleen, thymus, prostate, ovary, small intestine, colon or peripheral leukocyte (Figure 4.13A). The same blot hybridised to a *G3PDH* control probe (section 2.22) is shown in Figure 4.13B. Strong *G3PDH* bands were detected in all lanes.

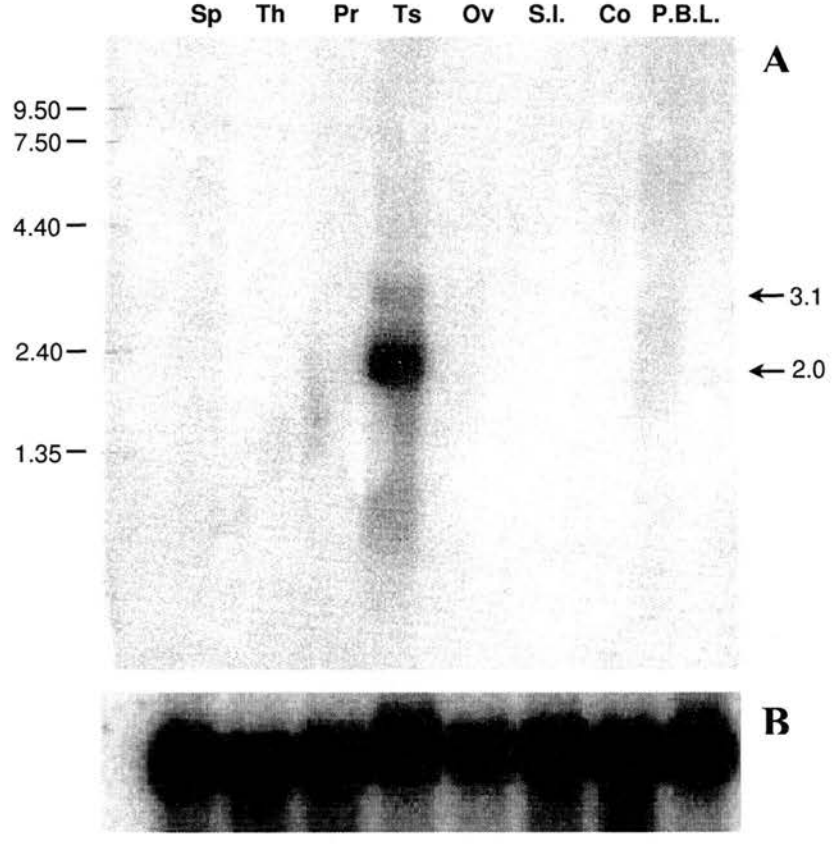


Figure 4.13: A commercial human multiple tissue northern blot (MTN2, Clontech) was probed with (**A**) a radiolabelled 3' *RPGRIP* probe and (**B**) a *G3PDH* probe. A strong band of 2.0 kb and a weak band of 3.1 kb were detected in testis only. Sp = spleen, Th = thymus, Pr = prostate, Ts = testis, Ov = ovary, S.I. = small intestine, Co = colon, P.B.L. = peripheral blood leukocyte. Sizes are in kilobases.

4.5.1.ii - Northern blot 2

A second commercial human multiple tissue northern blot (MTN, Clontech, Basingstoke) was probed with a 2759 bp ^{32}P-labelled *RPGRIP* fragment corresponding to RPGRIP nucleotides 80 to 2838 (encompassing exons 1 to 15, see section 2.22, page 69). A very weak 6.2 kb band was identified in the lane containing skeletal muscle RNA but no signal was obtained in other tissues, including pancreas, kidney, liver, lung, placenta, brain and heart (Figure 4.14A). Figure 4.14B shows that this blot also hybridised to a *G3PDH* control probe. Strong *G3PDH* bands were detected in skeletal muscle and heart but only weak bands in pancreas, kidney, liver, lung, placenta and brain.

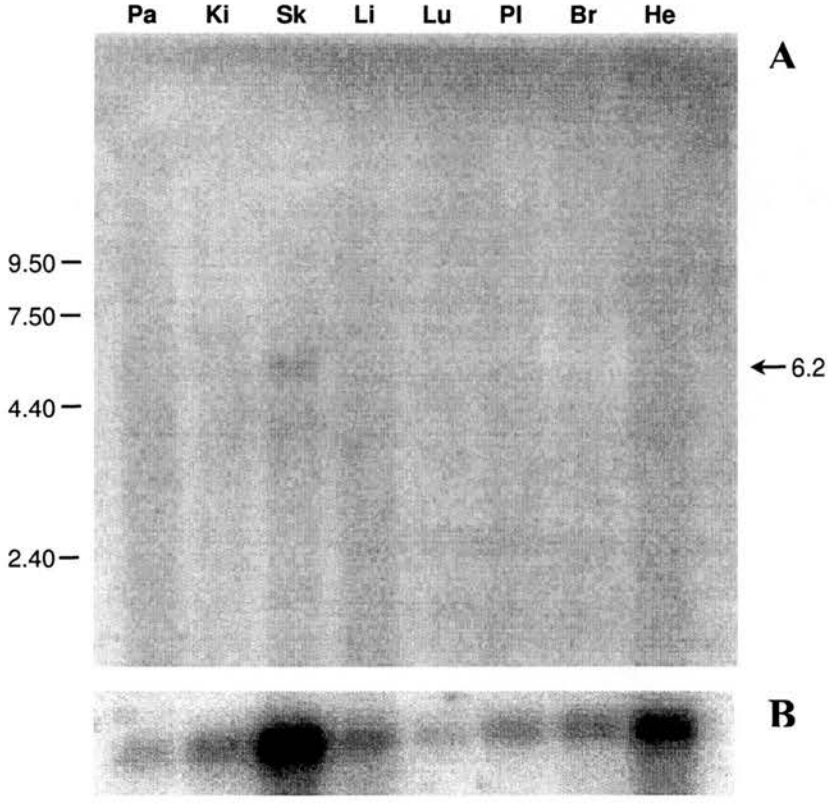


Figure 4.14: A commercial human multiple tissue northern blot (MTN, Clontech) was probed with (**A**) a radiolabelled 3' *RPGRIP* probe and (**B**) a *G3PDH* probe. A weak band of 6.2 kb was detected in skeletal muscle only. Pa = pancreas, Ki = kidney, Sk = skeletal muscle, Li = liver, Lu = lung, Pl = placenta, Br = brain, He = heart. Sizes are in kilobases.

4.5.1.iii - Northern blot 3

A northern blot was prepared using poly $(A)^+$ RNA purified from a range of human tissue samples including retina. Upon hybridisation to a 595 nucleotide radiolabelled *RPGRIP* probe corresponding to *RPGRIP* nucleotides 2206 to 2800 (from exon 11 to the end of exon 15, see section 2.22, page 69), a very faint signal was detected in the retina lane but no signal was seen in thymus, spleen, lung, liver, kidney, heart or brain. There appeared to be two bands, one of approximately 6.25 kb, and one of approximately 5 kb (Figure 4.15A). The blot was probed with ^{32}P-labelled β-*actin* (obtained by radiolabelling a commercial β-*actin* probe; Clontech, Basingstoke) and a strong band of the correct size was seen in retina only. Weak bands were seen in thymus, spleen, lung, heart and brain and no signal was detected in liver or kidney RNA (Figure 4.15B).

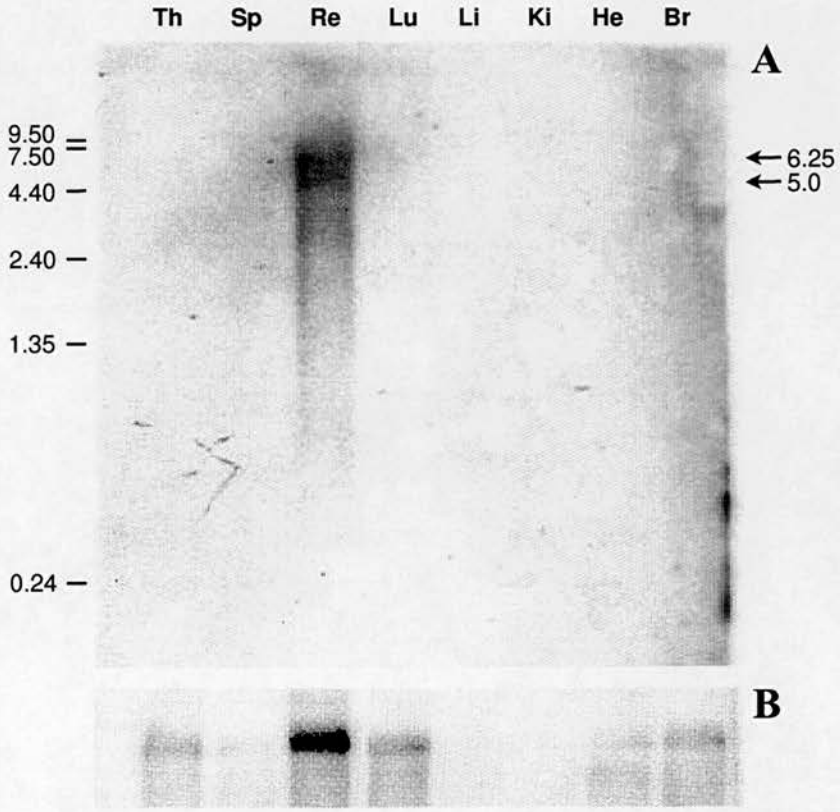


Figure 4.15: A Northern blot prepared using poly $(A)^+$ RNA purified from a range of human tissue samples was probed with (**A**) a radiolabelled 3' *RPGRIP* probe and (**B**) a β-*actin* probe. Two very faint bands, of 6.25 kb and 5.0 kb were detected in retina only. Th = thymus, Sp = spleen, Re = retina, Lu = lung, Li = liver, Ki = kidney, He = heart, Br = brain. Sizes are in kilobases.

Blots 2 and 3 failed to provide convincing results and it can be seen from the control probings (β-*actin* in the case of northern blot 3, *G3PDH* for northern blot 2) that they contain unequal loadings of RNA. On blot 2, a commercial product, there appears to be much more RNA in the skeletal muscle and heart lanes than the others and a very faint band was visible only in the skeletal muscle lane after probing with *RPGRIP*. Blot 3 contains much more RNA in the retina lane and, again, faint bands could be seen here only. From the control probing of northern blot 3 it appears that the spleen, liver and kidney lanes contain barely detectable amounts of RNA. The results from blots 2 and 3 suggest that with equal and sufficient amounts of RNA in each lane useful information about *RPGRIP* expression may have been collected.

4.5.2 - RT-PCR

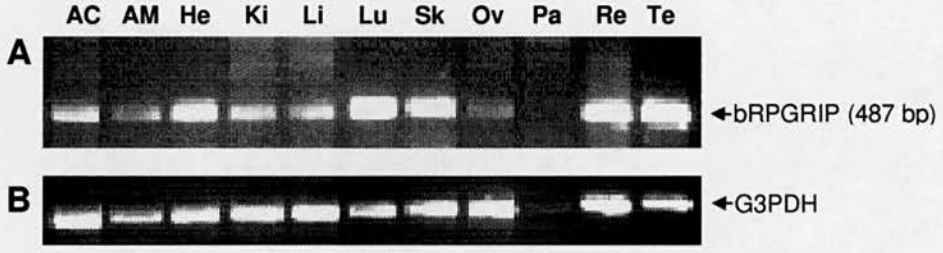


Figure 4.16: Amplification of bovine *RPGRIP* (A) and *G3PDH* (B) by RT–PCR from RNA isolated from a variety of tissues (AC, adrenal cortex; AM, adrenal medulla; He, heart; Ki, kidney; Li, liver; Lu, lung; Sk, skeletal muscle; Ov, ovary; Pa, pancreas; Re, retina; Te, testis). A band of the expected size (487 bp) was amplified from all tissues. The low signal from the pancreatic sample appears to result from poor quality RNA.

In order to carry out a more sensitive screen for low level *RPGRIP* expression, samples of human and bovine total RNA were used to amplify *RPGRIP* by RT-PCR. Bovine total RNA was obtained from Dr R. Vervoort (MRC Human Genetics Unit, Edinburgh.). These RNA samples were purified from tissues collected within 30 minutes of death and immediately homogenised in Trizol reagent (Gibco BRL) on ice. Samples were kept as Trizol homogenates at -70°C until they were used for RNA purification. The samples were treated in this way in order to protect the RNA from RNase degradation (Chomczynski and Sacchi, 1987). One round of RT-PCR-amplification (35 cycles) primed with bovine *RPGRIP* primers Bov FWD and Bov REV (Table 2.1, page 59) was sufficient to produce a band of the correct size from the majority of the samples (all except pancreas RNA, from which it

was only possible to produce a *G3PDH* product of very low abundance). Figure 4.16A shows the products of this reaction separated in an agarose gel. Figure 4.16B shows the results of a control reaction in which the same samples were used to amplify bovine *G3PDH* (using the BovG3PDHS and BovG3PDHAS PCR primers).

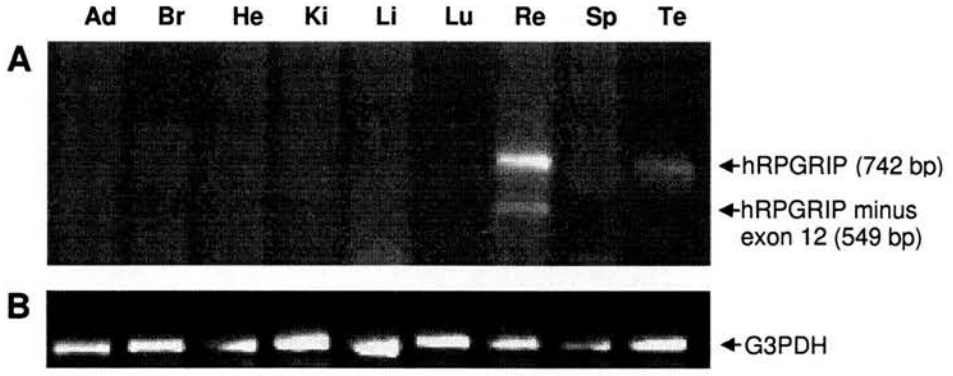


Figure 4.17: Amplification of human *RPGRIP* (A) and *G3PDH* (B) by RT-PCR using RNA from a variety of tissues (Ad, adrenal; Br, brain; He, heart; Ki, kidney; Li, liver; Lu, lung; Re, retina; Sp, spleen; Te, testis). A band of the expected size (742 bp) was amplified from retina and testis only. An additional smaller band (549 bp) was also amplified from testis. This band was sequenced and found to lack exon 12.

Human total RNA samples were obtained from previously frozen tissues collected from approximately 24 hours-old cadavers. Using the 1048 FWD and 544 REV oligonucleotide primers, which anneal to exons 9 and 14 respectively, *RPGRIP* transcripts were amplified from human retina and testis RNA after one 35-cycle round of amplification (Figure 4.17A). It was not possible to amplify *RPGRIP* from adrenal, brain, heart, kidney, liver, lung or spleen RNAs. A single product of the correct size (742 bp) was amplified from testis and retina RNA and an additional smaller product (549 bp) was amplified from retina RNA. The additional retina product (the lower retina band in Figure 4.17A) was sequenced and found to lack exon 12. The *G3PDH* positive control (amplified using the G3PDHF and G3PDHR PCR primers) is shown in Figure 4.17B. It was also possible to amplify *RPGRIP* exons 1 to 4 and 8 to 15 from human lymphoblastoid cell line derived total RNA (no attempt was made to amplify exons 5, 6 and 7) in the course of attempts to screen ARRP patients for *RPGRIP* mutations (see section 4.7, page 149). Specific products were only obtained following a second round of amplification using nested primers.

4.5.3 - Tissue *in situ* hybridisation

Sense and antisense DIG-labelled mouse RPGRIP riboprobes were prepared by cutting the aa204546 IMAGE EST clone with restriction endonucleases and incubating with T7 and T3 RNA polymerases respectively (see section 2.23). These were used, with the help of Judy Fletcher (MRC Human Genetics Unit, Edinburgh), in a tissue *in situ* hybridisation experiment with 12.5 and 15 day-old mouse embryo sections. No signal was detectable after a single experiment and unfortunately, due to time constraints, this was not repeated.

4.5.4 - *RPGRIP* ESTs

One other method of assessing *RPGRIP* expression was to determine the tissue source of *RPGRIP* ESTs identified in database searches. A large number of cDNA libraries have been sequenced and the ESTs deposited in databases, making this a valuable means of assessing gene expression. For example, since the commencement of the Washington University-Merck human EST project (Hillier *et al.*, 1996), 631,372 ESTs have been sequenced (http://genome.wustl.edu/est/esthmpg.html). *RPGRIP* ESTs came from a limited range of human tissues, namely foetal liver/spleen (r93221), retina (w28191, w26173), testis (ai015922, aa782963), mixed foetal lung, testis and B-cell (aa928161, aw081763) and pooled germ cell tumour (ai964059, ai655818, ai632512) (Table 4.5). r93221 was the only *RPGRIP* EST present in a library that did not have retina or testis origins (foetal liver/spleen). This indicates that *RPGRIP* is poorly represented in the EST databases, possibly as a result of a narrow range of expression, and, based on the number of clones from expressing tissues, the level of expression appears to be low.

Table 4.5: Tissue sources of human *RPGRIP* ESTs.

Library source	EST
Foetal liver/spleen	r93221
Retina	w28191, w26173
Testis	ai015922, aa782963
Mixed foetal lung, testis and b-cell	aa928161, aw081763
Pooled germ cell tumour	ai964059, ai655818, ai632512

4.6 - AMINO ACID SEQUENCE ANALYSIS

4.6.1 – Prediction of coiled-coil and C_2-domains

The deduced RPGRIP amino acid sequence was analysed using the HGMP-RC PIX assembly of bioinformatics programs (http://menu.hgmp.mrc.ac.uk/menu-bin/Pix) in an attempt to identify possible secondary or tertiary structural features. Such features were found by the Coils (Lupas, Van Dyke, and Stock, 1991), PROSITE (Hofmann *et al.*, 1999) and SOSUI (Hirokawa, Boon-Chieng and Mitaku, 1998) programs but the majority of the programs run during the PIX analysis failed to produce significant results for the RPGRIP sequence. These latter included prediction of subcellular localisation (Psort (Nakai and Horton, 1999)), BLAST searches against domain databases (SBASE (Murvai *et al.*, 2000), PRODOM (Corpet, Gouzy and Kahn, 1999), searches against motif and domain databases (Pfam (Bateman *et al.*, 2000), PRINTS (Atwood *et al.*, 2000), BLOCKS (Henikoff and Henikoff, 1991)), transmembrane predictions (Tmpred (Hofmann and Stoffel, 1993), Tmap (Persson and Argos, 1994), DAS (von Heijne, 1992), Phd (Rost and Sander, 1993)) and signal peptide predictions (Sigcleave (von Heijne, 1986)).

RPGRIP was predicted by the Coils program (Lupas, Van Dyke, and Stock, 1991) to have two coiled-coil domains at the N-terminus, the first encompassing residues 14 to 78, and the second encompassing residues 150 to 200 (Figures 4.11). Analysis using the PROSITE protein motif and domain database (Hofmann *et al.*, 1999) revealed the presence of four potential N-glycosylation sites, one potential cyclic adenosine monophosphate- (cAMP) and cGMP-dependent protein kinase phosphorylation site, nine potential protein kinase C phosphorylation sites, fourteen potential casein kinase II phosphorylation sites, two potential tyrosine kinase phosphorylation sites, four potential N-myristoylation sites and three potential amidation sites in the RPGRIP protein (Figure 4.18 and Table 4.6). RPGRIP is predicted to be a soluble protein by the SOSUI system, which analyses proteins according to the physicochemical properties of the amino acid sequence (Hirokawa, Boon-Chieng and Mitaku, 1998). The molecular weight of the 902 amino acid *RPGRIP* translation was predicted to be 97 kDa and that of the 586 amino acid translation of *RPGRIP* minus exons 5 to 7 was predicted to be 67 kDa. Their isoelectric points were predicted to be 4.99 and 4.93 respectively.

The RPGRIP (902 amino acid) sequence was also analysed using the Simple Modular Architecture Research Tool (SMART) internet-based database and research tool (Schultz *et al.*, 1998; Schultz *et al.*, 2000). This predicted two calcium- and phospholipid-

binding C_2-domains, encompassing residues 416-532 and 611-716 (Figures 4.11). Figure 4.18 shows a schematic drawing of the *RPGRIP* transcript, highlighting the positions of the predicted coiled-coil and C_2-domains in the primary sequence.

Table 4.6: Results of a search of the PROSITE database for potential motifs present in the 902 residue RPGRIP protein.

Protein Motifs	Position in RPGRIP
N-glycosylation sites	515-518; 523-526; 532-535; 652-655
cAMP- and cGMP-dependent protein kinase phosphorylation site	115-118
protein kinase C phosphorylation sites	314-316; 517-519; 566-568; 598-600; 619-621; 723-725; 731-733; 777-779; 855-857
casein kinase II phosphorylation sites	102-105; 127-130; 131-134; 161-164; 273-276; 446-449; 448-451; 475-478; 570-573; 571-574; 712-715; 770-773; 855-858; 868-871
tyrosine kinase phosphorylation sites	836-844; 837-844
N-myristoylation sites	193-198; 397-402; 499-504; 594-599; 703-708;
amidation sites	397-400; 623-626; 803-806

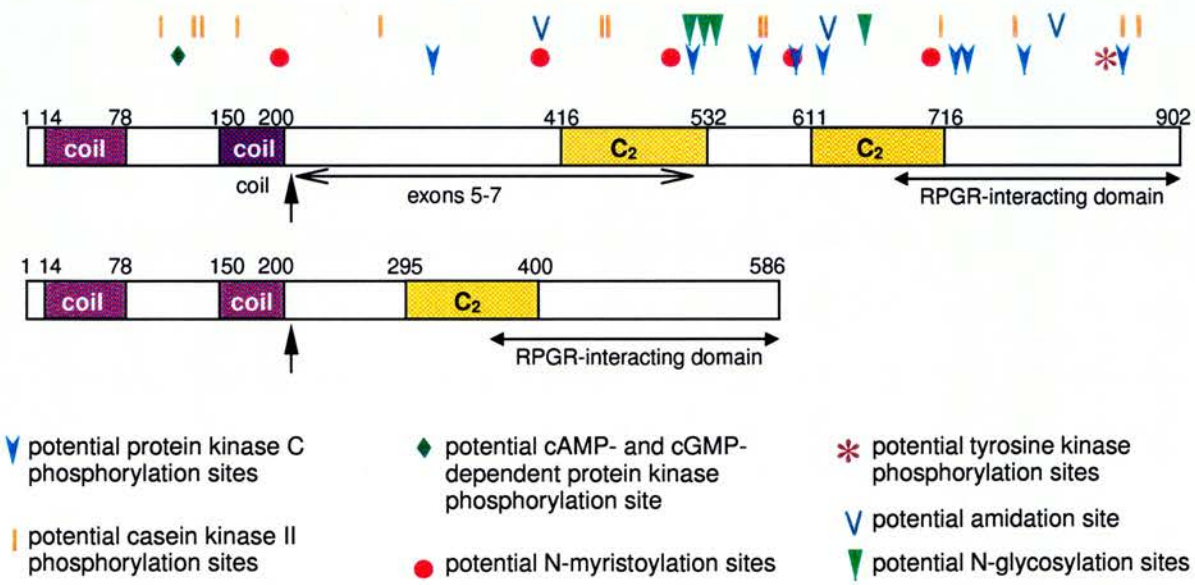


Figure 4.18: Schematic representations of the *RPGRIP* transcript with (upper) and without (lower) exons 5-7 (shaded). The two predicted coiled-coil encoding regions are shown in purple and the predicted C_2-domains are in yellow. The C-terminal 225 amino acids (labelled the "RPGR-interacting domain") were shown to interact with RPGR. The arrows show the position of the non-canonical splice site at the exon 4/exon 5 junction. Numbers indicate amino acid positions. Above the transcripts are shown the positions of potential protein motifs identified as a result of a PROSITE analysis (Hofmann *et al.*, 1999).

4.6.2 – Database searches for sequences with homology to RPGRIP

The available databases were used to search for proteins with homology to the translation of the full-length *RPGRIP* cDNA (encoded by exons 1-15). A number of sequences were identified from species including *Homo sapiens*, *Caenorhabditis elegans*, *Rattus norvegicus*, *Xenopus laevis*, *Mus musculus*, *Schizosaccharomyces pombe* and others (Table 4.7). The amino acid sequence encoded by the twelve upstream GENSCAN-predicted exons was then added to pRPGRIP and the resulting sequence was also used in BLAST database searches (Table 4.7). This extra sequence was also investigated as a single, separate polypeptide but there appeared to be no significant homologies. The majority of the sequences detected in database searches contained homology to RPGRIP in the region of the predicted coiled-coil domains (residues 14 to 78 and 150 to 200). To determine whether these predicted domains were masking any other (albeit weaker) homologies, residues 1-200 were removed from the RPGRIP sequence and more searches were performed (Table 4.7). The only sequences with homology to RPGRIP (minus the coiled-coil domains) were KIAA1005 protein (*Homo sapiens*) and 'weak similarity to myosin~cDNA' (*Caenorhabditis elegans*). The higher e values of the next most similar sequences indicate the lower significance of the subsequent hits.

KIAA1005 protein (*Homo sapiens*) and weak similarity to myosin~cDNA (*Caenorhabditis elegans*) were the two most significant matches on all three BLAST searches (with e values of $1x10^{-135}$ and $5x10^{-13}$ respectively). KIAA1005 is an uncharacterised human protein of 1,055 amino acids, which is 34% identical and 53% similar to human RPGRIP. It was identified during a project to deduce the coding sequence of previously unidentified human genes by sequencing a number of clones isolated from human brain cDNA libraries (Nagase *et al.*, 1999). The low e value of the RPGRIP/KIAA1005 alignment suggests a much greater significance than for the (uncharacterised) weak similarity to myosin~cDNA, which is 22% identical and 39% similar to human RPGRIP. An alignment of RPGRIP and KIAA1005 is shown in Figure 4.19.

The next eight sequences most similar to the 902 amino acid RPGRIP sequence (containing the predicted coiled-coil domains but lacking the upstream GENSCAN-predicted exons) were analysed using the Coils (coiled-coil prediction) program (Lupas, Van Dyke, and Stock, 1991). They were all found to contain domains predicted to adopt coiled-coil formations. The regions of homology coincide closely with the predicted RPGRIP coiled-coil domains and the BLAST search using the RPGRIP sequence without the coiled-coil domains failed to identify these proteins. These eight sequences consisted of three myosin

Table 4.7: Results of BLAST searches for sequences similar to RPGRIP (Hom. Reg. = homologous region, Ids = percentage of identical residues, +ves = percentage of similar residues, e value = the number of different alignments with equal or better scores expected to occur in a search by chance).

Sequence (accession number in brackets)	Hom. Reg. (aa)	Ids (%)*	+ves (%)**	e value	Score
*A. **RPGRIP sequence (amino acids 1-902, equivalent to exons 1-15)***					
1. KIAA1005 protein (AB023222, *Homo sapiens*)	35-900	34	53	$1x10^{-135}$	484
2. Weak similarity to myosin~cDNA (Z46791, *Caenorhabditis elegans*)	50-555	22	39	$5x10^{-13}$	78
3. Hypothetical 130.5 kDa protein (C09G5.8, *Caenorhabditis elegans*)	50-398	21	37	$1x10^{-7}$	60
4. 364K Golgi complex-associated protein (GCP360, *Rattus norvegicus*)	34-261; 45-196; 17-192; 35-211	21; 23; 20; 18	45; 44; 39; 44	$3x10^{-4}$; 1.4; 4.1; 7.0	48
5. Myosin of the dilute-myosin-V family (U60416, *Rattus norvegicus*)	49-226	24	47	0.002	46
6. Lamin A (S02358, *Xenopus laevis*)	34-195	26	47	0.004	45
7. Lamin B2 (X54098, *Mus musculus*)	34-199	23	46	0.004	45
8. Myosin-3 isoform, heavy chain (Z98762, *Schizosaccharomyces pombe*)	38-202	26	45	0.006	44
9. Hook1 protein (AF044923, *Homo sapiens*)	44-192	28	47	0.006	44
10. Myosin-II; Myp2p (AF029788, *Schizosaccharomyces pombe*)	38-202	26	45	0.006	44
*B. **RPGRIP sequence including the 12 upstream GENSCAN predicted exons (amino acids 1-1431)***					
1. KIAA1005 protein (AB023222, *Homo sapiens*)	355-1429	33	52	$1x10^{-159}$	562
2. Weak similarity to myosin~cDNA (Z46791, *Caenorhabditis elegans*)	348-1084	22	39	$3x10^{-16}$	89
3. Hook1 protein (AF044923, *Homo sapiens*)	362-721	23	37	$3x10^{-8}$	62
4. Integrin homolog (S30782, *Saccharomyces cerevisiae*)	414-724; 403-726; 366-702; 404-738	22; 23; 19; 17	43; 44; 39; 38	$5x10^{-8}$; $9x10^{-4}$; 0.008; 0.15	62
5. Integrin analogue gene /Intracellular transport protein USO1 (X54378, *Saccharomyces cerevisiae*)	414-724; 365-762; 404-738	21; 19; 17	42; 40; 38	$9x10^{-8}$; 0.088; 0.15	61
6. Cingulin (AF207901, *Xenopus laevis*)	367-728; 428-728; 375-738	19; 22; 21	41; 43; 41	$3x10^{-7}$; $2x10^{-4}$; 0.013	59
7. Unknown protein (AC011665, *Arabidopsis thaliana*)	367-710	20	42	$4x10^{-7}$	59
8. Unnamed protein product (AK025690, *Homo sapiens*)	410-721; 414-705	23; 20	41; 35	$6 x10^{-7}$; 0.15	58
9. Similar to nuclear matrix constituent protein 1 (AC007357, *Arabidopsis thaliana*)	404-747; 389-737	23; 20	39; 38	$6 x10^{-7}$; $2x10^{-4}$	58

10. Paramyosin (S04027, *Caenorhabditis elegans*)	417-779; 367-728	21; 19	38; 38	1×10^{-6}; 0.003	57
*C. **RPGRIP sequence minus the coiled-coil domains (amino acids 201-902)***					
1. KIAA1005 protein (AB023222, *Homo sapiens*)	205-900	34	53	1×10^{-113}	408
2. Weak similarity to myosin~cDNA (Z46791, *Caenorhabditis elegans*)	269-555	25	42	8×10^{-8}	60
3. Unknown (AF115283, *Leptospira interrogans*)	653-786	29	47	0.008	43
4. Hypothetical 130.5 kDa protein (C09G5.8, *Caenorhabditis elegans*)	269-398	26	42	0.015	43
5. Nlj21 (AF000402, *Lotus japonicus*)	674-845	23	42	0.096	40
6. Hypothetical protein L2719.04 (AL117324, *Leishmania major*)	239-332	28	44	0.16	39

proteins (from *Rattus norvegicus* and *Schizosaccharomyces pombe*), two proteins of unknown function (hypothetical 130.5 kDa protein of *Caenorhabditis elegans* and 364K Golgi complex-associated protein of *Rattus norvegicus*), two lamins (lamin A, *Xenopus laevis* and lamin B2, *Mus musculus*) and human Hook1 protein. Lamins are members of the intermediate filament protein family that make up the fibrous lattice lying beneath the inner nuclear membrane (Stick 1992). Three types of mammalian lamins exist, A, B and C. Lamin B2 was identified as the second major B-type lamin after lamin B1 (Hoger *et al.*, 1990). The relatively high e value for the lamin A and B2 alignments (0.004) indicates that these matches are less likely to be significant than KIA10055. *Hook1* encodes a human homologue of the *Drosophila Hook* gene involved in endocytic trafficking, so-called because flies with hook mutations display a characteristic hooked-bristle phenotype (Kramer and Phistry, 1999). *Hook* mutations have a detrimental effect on the accumulation of soluble and transmembrane ligands into multivesicular bodies.

The e value for Hook1 is relatively high when aligned with the 902-amino acid RPGRIP protein (0.006) but this diminishes (to 3×10^{-8}) with the addition of the 529-amino acid translation of the additional upstream GENSCAN predicted exons. The homology of Hook1 to the former polypeptide (the 902 amino acid RPGRIP sequence) encompassed the predicted coiled-coil domains. This region was extended towards the N-terminus in the case of the 1431-residue protein (the coiled-coil domains occupy amino acids 543-607 and 679-729 of the 1431 residue RPGRIP sequence that includes the translation of the GENSCAN-predicted exons). Among the remaining hits with the highest score/lowest e value for this

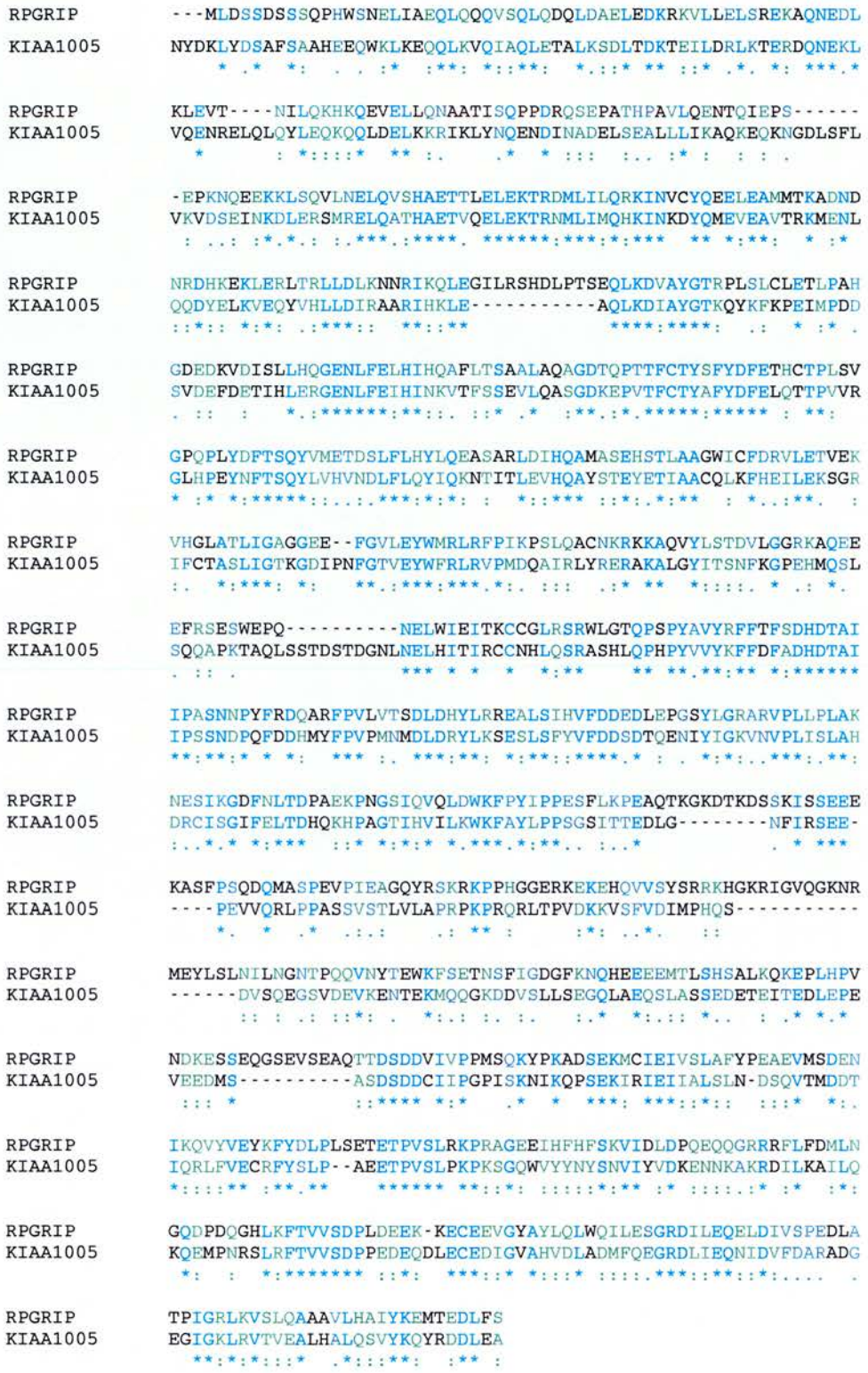


Figure 4.19: CLUSTAL W alignment (Thompson *et al.*, 1994) of the human RPGRIP and KIAA1005 proteins (* = single, fully conserved residue,: = conservation of strong groups, . = conservation of weak groups, blank space = no consensus).

larger hypothetical sequence were integrin homolog (*Saccharomyces cerevisiae*), intracellular transport protein USO1 (*Saccharomyces cerevisiae*), cingulin (*Xenopus laevis*), unknown protein (*Arabidopsis thaliana*), unnamed protein product (*Homo sapiens*) and nuclear matrix constituent protein 1 (*Arabidopsis thaliana*). Integrin homolog and uso 1 are homologous proteins. Uso 1 has been proposed to be a component of the cytoskeleton with a role in the transport of proteins from the ER to downstream secretory compartments (Nakajima *et al.*, 1991). It was identified from a *Saccharomyces cerevisiae* mutant strain blocked in the protein secretion pathway. At the C-terminal end of the 1,790 amino acid Uso1 protein is an alpha-helical structure of 1,100 residues similar to the coiled-coil region found in the cytoskeleton-related proteins (Nakajima *et al.*, 1991). Cingulin is a phosphoprotein found on the cytoplasmic face of epithelial tight junctions (Cordenonsi *et al.*, 1999). It forms coiled-coil parallel dimers and interacts with a number of proteins including myosin. It has been proposed to form a connection between the tight junction submembrane plaque domain and the actomyosin cytoskeleton. Nuclear matrix constituent protein 1 (NMCP1) was identified from carrot cells (*Daucus carota*) as a constituent of higher plant structures architecturally homologous to the nuclear lamina of vertebrates (Masuda *et al.*, 1997). The primary structure of NMCP1 is similar to myosin, tropomyosin, and some intermediate filament proteins (of which lamins A and B are examples – see Table 4.7). Its function is largely undetermined.

4.6.3 – Bovine and mouse RPGRIP

The RPGRIP amino acid sequences reported by Roepman *et al.* (2000) and Hong *et al.* (2000) were compared to the RPGRIP sequence described here. Hong *et al.* (2000) submitted a 1331 residue murine RPGRIP sequence to GenBank (accession number AY008297) and Roepman *et al.* (2000) submitted several sequences from bovine and human, the longest being a 1221 residue bovine sequence (accession number AAG10247). The longest human sequence reported by Roepman *et al.* (2000) (accession number AAG10246) contained 762 residues, 140 amino acids shorter than the 902 residue sequence described here. Neither Hong *et al.* (2000) nor Roepman *et al.* (2000) provide any information regarding the genomic structure of RPGRIP in human, bovine or mouse. Compared to the hRPGRIP sequence shown in Figure 4.11, the bovine sequence contained an additional 336 amino acids at the amino terminal end and the murine sequence an additional 345 amino acids. The 529 amino acids encoded by the (undetected) 12 upstream GENSCAN-predicted exons (exons -12 to -1) were added to the 902 encoded by the detected exons (exons 1 to 15) and the resulting 1431 aa sequence was compared to the 1221 aa

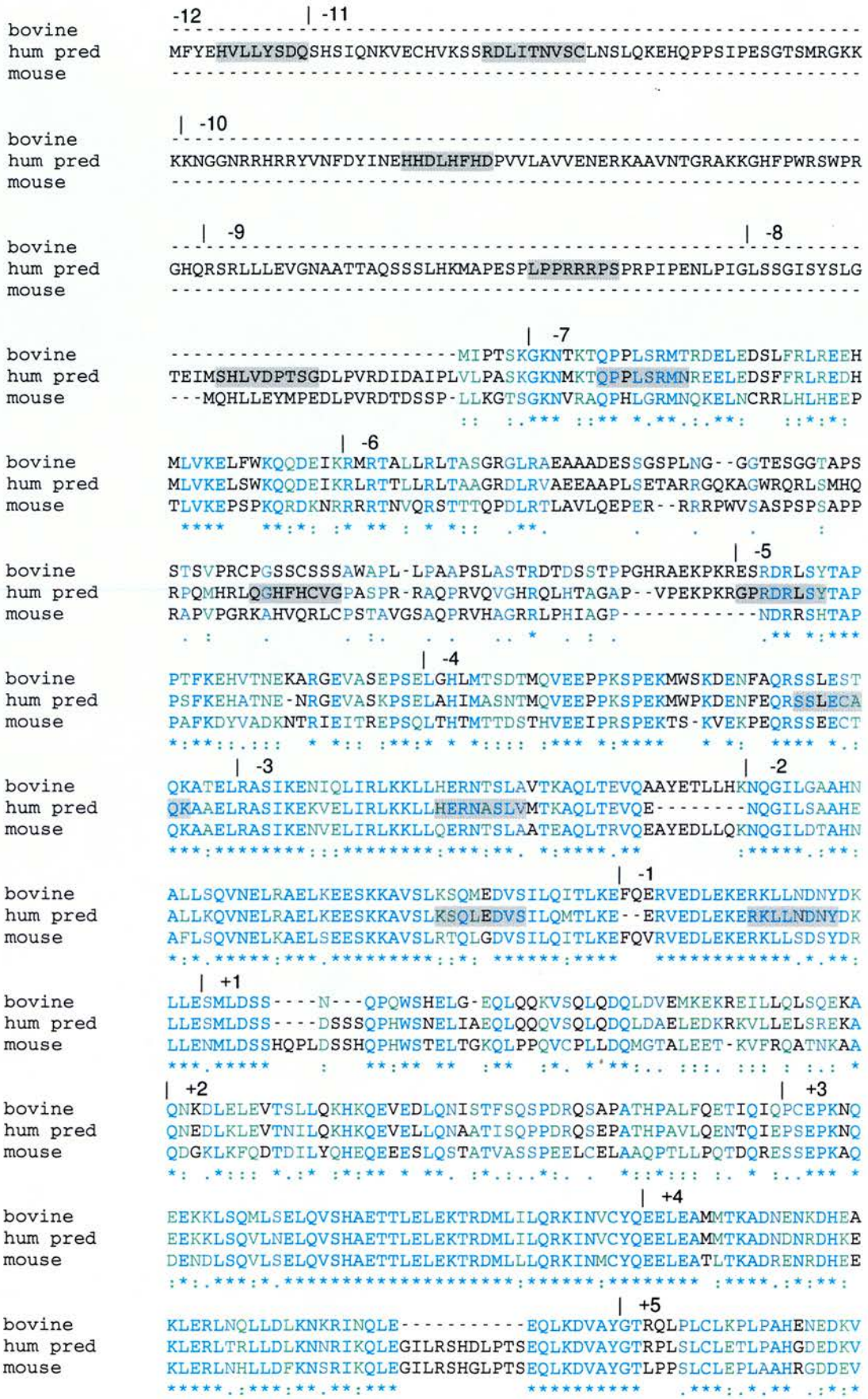
bovine and 1331 aa murine sequences. Residues 1 to 183 of this predicted hRPGRIP were found to have no equivalent in the bovine and murine proteins but the remaining sequence (residues 184 to 1431) was found to be similar to both over their entire lengths (the hypothetical human protein was 73% identical and 81% similar to the bovine and 59% identical and 68% similar to the mouse). These three sequences are shown in alignment in Figure 4.20. Although none of the 12 upstream GENSCAN-predicted exons appeared to be transcribed in retina (Figure 4.8), the close similarity of the deduced amino acid sequence encoded by exons -8 to -1 to bovine and mouse RPGRIP sequences suggests that they may in fact be part of the human *RPGRIP* gene. The initiation codon of the mouse sequence coincides with a methionine codon in the hypothetical human sequence. It is possible that this is the initiation codon of an RPGRIP transcript. The DNA surrounding this codon in comparison with the consensus sequence for initiation of translation (Kozak 1987) is shown in Table 4.8.

Table 4.8: The sequence around a putative translation start site of the hypothetical *hRPGRIP* sequence compared with the consensus determined by Kozak (1987). The numbers indicate the percentage of transcripts that have the residue shown in that position.

Position:	-9	-8	-7	-6	-5	-4	-3	-2	-1	+1	+2	+3	+4
Consensus:	G^{33}	C^{39}	C^{37}	G^{44}	C^{39}	C^{53}	A^{61}/G^{36}	C^{49}	C^{55}	A	T	G	G^{46}
***hRPGRIP*:**	A^{23}	C	A^{23}	G	A^{18}	G^{15}	A	T^{11}	C	A	T	G	T^{15}

There are two obvious gaps in the hypothetical human RPGRIP sequence which are absent from the bovine and mouse. The first of these is an 8 residue gap in the region encoded by the 3' end of exon -3. By adding the 4 nucleotides of genomic sequence immediately downstream of exon -3 and 20 from immediately upstream of exon -2 (see Table 4.3), it is possible to fill this gap and obtain canonical splice sites. This produces the following amino acid sequence (additional sequence in red):

```
                                        -3  |  -2
bovine  ...T E V Q A A Y E T L L H K |N Q G I...
human   ...T E V Q E V S T F L L T Q |N Q G I...
mouse   ...T R V Q E A Y E D L L Q K |N Q G I...
```


-12
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
+1
+2
+3
+4
+5
bovine
hum pred
mouse

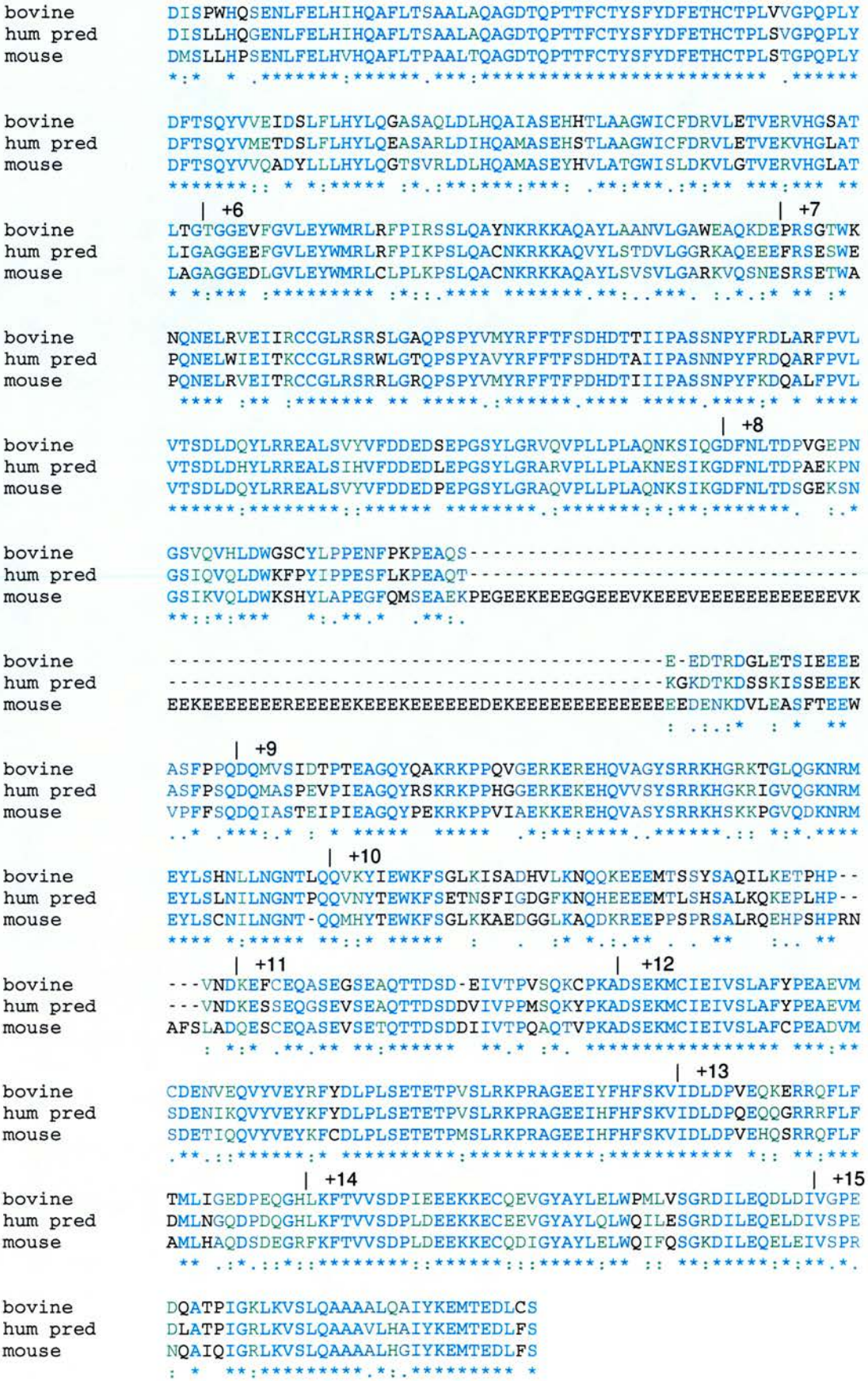

bovine DISPWHQSENLFELHIHQAFLTSAALAQAGDTQPTTFCTYSFYDFETHCTPLVVGPQPLY
hum pred DISLLHQGENLFELHIHQAFLTSAALAQAGDTQPTTFCTYSFYDFETHCTPLSVGPQPLY
mouse DMSLLHPSENLFELHVHQAFLTPAALTQAGDTQPTTFCTYSFYDFETHCTPLSTGPQPLY
bovine DFTSQYVVEIDSLFLHYLQGASAQLDLHQAIASEHHTLAAGWICFDRVLETVERVHGSAT
hum pred DFTSQYVMETDSLFLHYLQEASARLDIHQAMASEHSTLAAGWICFDRVLETVEKVHGLAT
mouse DFTSQYVVQADYLLLHYLQGTSVRLDLHQAMASEYHVLATGWISLDKVLGTVERVHGLAT
| +6
| +7
bovine LTGTGGEVFGVLEYWMRLRFPIRSSLQAYNKRKKAQAYLAANVLGAWEAQKDEPRSGTWK
hum pred LIGAGGEEFGVLEYWMRLRFPIKPSLQACNKRKKAQVYLSTDVLGGRKAQEEEFRSESWE
mouse LAGAGGEDLGVLEYWMRLCLPLKPSLQACNKRKKAQAYLSVSVLGARKVQSNESRSETWA
bovine NQNELRVEIIRCCGLRSRSLGAQPSPYVMYRFFTFSDHDTTIIPASSNPYFRDLARFPVL
hum pred PQNELWIEITKCCGLRSRWLGTQPSPYAVYRFFTFSDHDTAIIPASNNPYFRDQARFPVL
mouse PQNELRVEITRCCGLRSRRLGRQPSPYVMYRFFTFPDHDTIIIPASSNPYFKDQALFPVL
| +8
bovine VTSDLDQYLRREALSVYVFDDEDSEPGSYLGRVQVPLLPLAQNKSIQGDFNLTDPVGEPN
hum pred VTSDLDHYLRREALSIHVFDDEDLEPGSYLGRARVPLLPLAKNESIKGDFNLTDPAEKPN
mouse VTSDLDQYLRREALSVYVFDDEDPEPGSYLGRAQVPLLPLAQNKSIKGDFNLTDSGEKSN
bovine GSVQVHLDWGSCYLPPENFPKPEAQS----------
hum pred GSIQVQLDWKFPYIPPESFLKPEAQT----------
mouse GSIKVQLDWKSHYLAPEGFQMSEAEKPEGEEKEEEGGEEEVKEEEVEEEEEEEEEEEVK
bovine -----------E-EDTRDGLETSIEEEE
hum pred -----------KGKDTKDSSKISSEEEK
mouse EEKEEEEEEEREEEEEKEEEKEEEEEDEKEEEEEEEEEEEEEEDENKDVLEASFTEEW
| +9
bovine ASFPPQDQMVSIDTPTEAGQYQAKRKPPQVGERKEREHQVAGYSRRKHGRKTGLQGKNRM
hum pred ASFPSQDQMASPEVPIEAGQYRSKRKPPHGGERKEKEHQVVSYSRRKHGKRIGVQGKNRM
mouse VPFFSQDQIASTEIPIEAGQYPEKRKPPVIAEKKEREHQVASYSRRKHSKKPGVQDKNRM
| +10
bovine EYLSHNLLNGNTLQQVKYIEWKFSGLKISADHVLKNQQKEEEMTSSYSAQILKETPHP--
hum pred EYLSLNILNGNTPQQVNYTEWKFSETNSFIGDGFKNQHEEEEMTLSHSALKQKEPLHP--
mouse EYLSCNILNGNT-QQMHYTEWKFSGLKKAEDGGLKAQDKREEPPSPRSALRQEHPSHPRN
| +11
| +12
bovine ---VNDKEFCEQASEGSEAQTTDSD-EIVTPVSQKCPKADSEKMCIEIVSLAFYPEAEVM
hum pred ---VNDKESSEQGSEVSEAQTTDSDDVIVPPMSQKYPKADSEKMCIEIVSLAFYPEAEVM
mouse AFSLADQESCEQASEVSETQTTDSDDIIVTPQAQTVPKADSEKMCIEIVSLAFCPEADVM
| +13
bovine CDENVEQVYVEYRFYDLPLSETETPVSLRKPRAGEEIYFHFSKVIDLDPVEQKERRQFLF
hum pred SDENIKQVYVEYKFYDLPLSETETPVSLRKPRAGEEIHFHFSKVIDLDPQEQQGRRRFLF
mouse SDETIQQVYVEYKFCDLPLSETETPMSLRKPRAGEEIHFHFSKVIDLDPVEHQSRRQFLF
| +14
| +15
bovine TMLIGEDPEQGHLKFTVVSDPIEEEKKECQEVGYAYLELWPMLVSGRDILEQDLDIVGPE
hum pred DMLNGQDPDQGHLKFTVVSDPLDEEKKECEEVGYAYLQLWQILESGRDILEQELDIVSPE
mouse AMLHAQDSDEGRFKFTVVSDPLDEEKKECQDIGYAYLELWQIFQSGKDILEQELEIVSPR
bovine DQATPIGKLKVSLQAAAALQAIYKEMTEDLCS
hum pred DLATPIGRLKVSLQAAAVLHAIYKEMTEDLFS
mouse NQAIQIGRLKVSLQAAAALHGIYKEMTEDLFS

Figure 4.20 (previous 2 pages): CLUSTAL W alignment (Thompson *et al.*, 1994) of bovine, mouse and hypothetical human (hum pred) RPGRIP amino acid sequences. The human sequence is a translation of the gene predicted by analysing the human chromosome 14 genomic sequence (accession number AL135744) with the GENSCAN program (Burge and Karlin 1998). The numbers and lines correspond to human exons. Shaded letters indicate the residues encoded by DNA corresponding to the target sites of the PCR primers used to test for the presence of these exons (section 4.3). * = single, fully conserved residue, : = conservation of strong groups, . = conservation of weak groups, blank space = no consensus.

The second gap is in the region encoded by the 5' end of exon -1. The addition to exon -1 of 6 nucleotides of chromosome 14 genomic sequence from immediately upstream of exon -1 fills the gap and again provides a canonical splice acceptor site. These extra nucleotides would translate into phenylalanine and glutamine, identical to the mouse and bovine residues at these positions. This 1259 amino acid hypothetical human sequence (including the additional residues to fill the aforementioned gaps) was analysed using the Compute pI/Mw tool (Wilkins *et al.*, 1998) and predicted to have a molecular weight of 144 kDa and an isoelectric point of 5.47.

4.7 – SCREENING *RPGRIP* FOR MUTATIONS IN ARRP PATIENTS

Bruford (1996) attempted to find ARRP loci in the Sardinian population by studying 11 south-central Sardinian ARRP families, consisting of 28 affected and 44 unaffected individuals. This was done by screening the whole genome for linkage with 195 microsatellite markers. No linkage was detected to ARRP in the entire set of families (no lod scores > 3), although heterogeneity analysis (HOMOG program) showed suggestive linkage (P=0.065) in a subset of 5 families to chromosomal region 14q11. Examination of additional microsatellite markers in one of these families suggested that an ARRP gene may lie distal to marker D14S275. The location of the *RPGRIP* gene was estimated by radiation hybrid mapping using the Genebridge 4 panel (Gyapay *et al.*, 1996; see section 4.4.3, page 131) and was placed proximal to D14S275, between D14S264 and D14S275 on the framework of markers, in chromosomal region 14q11. *RPGRIP* was placed 6.94 cR_{6000} from the closest marker, D14S264, and the lod score of 2.03 means that this position is 107 times more likely to be the true position than the next most likely placement. Although Bruford concluded that linkage had been detected to markers distal to D14S275, the coincidence of finding *RPGRIP* localised to chromosome 14q11, and the comparatively low radiation hybrid score that placed it proximal to D14S275 (in addition to the possibility that the maker order is incorrect), led us to believe that it would be unwise to dismiss *RPGRIP* as

a candidate RP gene. Consequently, the screening of *RPGRIP* for mutations in Sardinian ARRP patients was begun.

A single individual was screened from three of the five Sardinian ARRP families that were identified as linked to 14q. Time constraints prevented investigation of the other two families. *RPGRIP* was amplified from RNA rather than directly from genomic DNA because, at the time this work was carried out, the complete structure of the *RPGRIP* gene had not yet been determined. In addition, this was prior to the identification of exons 5-7, so these were not screened. Frozen lymphoblastoid cell lines from the three affected individuals (BB13 from the ARRP4 pedigree, BB15 from ARRP1, and BB39 from ARRP12) were thawed and re-established by Alan Lennon (MRC Human Genetics Unit). Total RNA was purified from cell pellets and used to amplify *RPGRIP* cDNA by RT-PCR as outlined in section 2.30 (page 82). A number of sequence variations from wild-type *RPGRIP* were found in the cDNAs from each patient. These were then re-examined by sequencing *RPGRIP* PCR products containing the corresponding genomic DNA fragments (section 2.30, page 81). At the genomic level, the sequence of *RPGRIP* in these three patients (coding regions and splice sites) was identical to the wild-type *RPGRIP* cDNA. This indicated that the changes seen at the cDNA level were likely to have resulted from reverse-transcriptase errors rather than mutations.

4.8 – DISCUSSION

The first aim of the experiments presented in this chapter was to isolate the full length *RPGRIP* cDNA. Primer extension methods were used to obtain a 2894 bp *RPGRIP* cDNA. Initial attempts to extend the *RPGRIP* cDNA (obtained from human EST r93221) using 5' RACE were unsuccessful so primer extension methods were applied to a library of adaptor-ligated linear genomic DNA fragments. In this way, the *RPGRIP* open reading frame was extended by 146 bp in the 5' direction. This new sequence was then extended at the cDNA level in human retina and testis 5' RACE reactions. The retina and testis 5' RACE products both contain in-frame stop codons upstream of the putative initiation codon. This would suggest that the exons starting these transcripts are complete at the 5' end. The sequence surrounding the presumed initiation codon conforms well to the consensus identified by Kozak (1987), reinforcing the argument that this codon is used to start translation *in vivo*. A TATA box has not been identified upstream of the first methionine triplet. The promoters of many genes lack such elements (Novina and Ror, 1996; Audic and Claverie, 1998) including many of the so-called 'housekeeping' genes (Azizkhan *et al.*, 1993). GC and CAAT boxes were also absent from the sequence immediately upstream of exon 1. These elements were

not found in the sequence upstream of GENSCAN-predicted exon -8. Further work is necessary to determine the promoter of *RPGRIP*.

CpG islands are segments of DNA closely associated with all housekeeping genes and many tissue-specific genes which are enriched in the CpG dinucleotide (reviewed by Kundu and Rao, 1999). They are defined as sequences over 200 bp in length (usually 1-2 kb) which contain at least 50% dGTP/dCTP with a CpG content greater than 0.6 of that expected on the basis of the dGTP/dCTP content of the surrounding DNA (Gardiner-Garden and Frommer, 1987). If ubiquitous *RPGRIP* expression had been detected, suggesting a housekeeping gene, positive identification of a CpG island may have helped to locate the 5' end of the gene. However, from the expression studies carried out it would appear that human *RPGRIP* is not a housekeeping gene. Although CpG islands are also found in some tissue-specific genes, none were predicted to be located close to *RPGRIP* exons 1 to 15. Potential CpG islands (of 699, 357, 259, 265, 288 and 1061 bp) were identified in close approximation with GENSCAN-predicted exons -9 and -6 following a GrailEXP (Hyatt *et al.*, 2000) analysis of the chromosome 14 genomic sequence containing *RPGRIP*.

3' RACE experiments carried out using testis cDNA identified a transcript extending 84 bp downstream from the open reading frame. Most of the ESTs identified in BLAST searches for sequences with homology to *RPGRIP* are from the 3' end of the cDNA and include the stop codon (the 903rd codon of *RPGRIP*). None of the ESTs or the 3' RACE product contain either of the two most common polyadenylation signals, AATAAA and ATTAAA (Graber *et al.*, 1999). The nearest AATAAA is found 5 kb 3' to the end of the *RPGRIP* open reading frame but the 3.1 kb transcript size seen in testis (Figure 4.13) is too small to include such a long 3' UTR. In the absence of a strong polyadenylation signal (AATAAA or ATTAAA) (Sheets *et al.*, 1990) it is likely that cleavage and polyadenylation of the mRNA transcripts may be directed by variant polyadenylation signals. Sheets *et al.* (1990) analysed 296 ESTs *in silico* and calculated the frequency of alternative residues at each position within the polyadenylation hexamers, as shown below.

	A	A	T	A	A	A
A	**98**	**86**	0.8	**98**	**95**	**96**
T	0.8	12	**98**	0.5	3.4	1.5
G	1.1	1.9	0.4	1.5	0	0.7
C	0.4	0.7	0.8	0.4	1.5	2.3

Graber *et al.*, (1999) looked at a much larger number of 3'-ESTs (4,427) and found that 35% of them did not have AATAAA or ATTAAA. The hexamers TATAAA and AGTAAA are found close to the 3' end of most *RPGRIP* transcripts (RACE products and ESTs) (see

Figure 4.10). According to Sheets *et al.* (1990), the polyadenylation signal is typically located 15-25 nucleotides upstream of the polyadenylation site. The last adenosine residue of the AGTAAA hexamer is 31 bp from the end of the *RPGRIP* transcripts and the presumed site of polyadenylation; the last adenosine residue of the TATAAA hexamer is 17 bp from the end of the transcripts. Graber et al (1999) estimated that TATAAA is the fourth most common of 18 single-base-substitution variants of AATAAA in 3'-ESTs that do not contain AATAAA. The variants seen more frequently than TATAAA are (in order of frequency) ATTAAA, AAAAAA and AGTAAA. Sheets *et al.* (1990) assessed the effects of point mutations in the AATAAA polyadenylation signal on the efficiency of polyadenylation *in vitro*. When the polyadenylation signal was AGTAAA, the efficiency of polyadenylation was 29% (± 8.1%) that of AATAAA. When the polyadenylation signal was TATAAA the efficiency of polyadenylation was 17% (± 3%) that of AATAAA. Thus, AGTAAA would appear to be a more efficient polyadenylation signal, at least *in vitro*, although TATAAA is closer to the supposed polyadenylation site. Both are relatively common variants of AATAAA (Graber *et al.*, 1999) and either could be acting as the polyadenylation signal, suggesting that the 3' end of the identified *RPGRIP* sequence does represent the correct 3' end.

Eight out of the nine human ESTs located at the 3' end of *RPGRIP* (ai964059, ai655818, ai632512, ai05922, aa782963, aw681763, aa928161 and w26173) terminate 81-83 bp downstream of the termination codon. These ESTs were oligo (dT)-primed, indicating (if the polyadenylation signal is located 15-25 nucleotides upstream of the polyadenylation site) that the polyadenylation signal is likely to be 57 - 67 bp downstream of the stop codon (the first position of the TATAAA hexamer is 59 bp downstream of the stop codon and the AGTAAA hexamer is 45 bp downstream). The presence of in-frame stop codons in these ESTs (at the position shown in Figure 4.10) also indicates that this is the likely 3' end of the coding sequence. The remaining EST (aa476670) is also oligo (dT)-primed but ends 315 bp 3' to the end of the termination codon. Immediately 3' to the end of this EST is an A-rich region (17/20), suggesting that an extended transcript has false-primed within this region.

In addition, probing the first northern blot (MTN2) revealed the presence of two transcripts in testis, of 3.1 kb and 2.0 kb. These sizes are close to the full length *RPGRIP* cDNA (i.e. including exons 5-7) and the smaller transcript (lacking exons 5-7), at 2.89 and 1.95 kb respectively. Estimates of transcript sizes from northern blots must be treated as rough indicators only (since they are made by comparing fragment movements with those of a limited number of standards), so the possibility of additional small 5' exons not revealed by RACE cannot be excluded.

A second goal of these experiments was to determine the genomic structure of *RPGRIP* and was facilitated by the appearance of chromosome 14 sequence in genomic sequence databases. This submission also enabled exon-finding analyses to be carried out, using the HGMP RC NIX collection of bioinformatic programs. The GENSCAN analysis (Burge and Karlin, 1998) predicted the existence of fifteen additional exons within the *RPGRIP* gene: twelve upstream of the first exon, and three between the fourth and fifth detected exons. The latter three exons were found to be transcribed in human retina but the twelve 5' predicted exons were not detected. The existence of additional exons had been proposed (i) on the basis of northern blot data (two bands were observed in testis (MTN2, Figure 4.13, page 133): a smaller band of 2.0 kb, roughly the size of the presumed full length RPGRIP cDNA, and a larger band of 3.1 kb) and (ii) the presence of an approximately 488 bp larger but less abundant band accompanying the expected 1862 bp product of a PCR primed in the 5' and 3' UTRs of *RPGRIP* (Figure 4.9)

It remains a possibility that the undetected 12 upstream exons (exons -12 to -1) are transcribed into an isoform that lacks exon 10 (containing the RG32longREV target sequence), or that a transcript containing these exons was not abundant in the retina cDNA which was used in the attempt to amplify these exons. Alternatively, the reactions carried out to amplify *RPGRIP* cDNA between the predicted exons and exon 10 may have failed due to the size limitations of RT-PCR. Comparison of a translation of the 12 upstream GENSCAN-predicted exons with recently reported bovine (Roepman *et al.*, 2000) and mouse (Hong *et al.*, 2000) RPGRIP proteins indicates that it is likely that exons -8 to -1 contribute to the *RPGRIP* gene. Further cDNA work using alternative primers and different tissue cDNAs is needed to conclusively include or exclude the presence of these predicted exons.

A third goal was to identify the chromosomal location of the *RPGRIP* gene. Screening a human monochromosomal somatic cell hybrid DNA panel (Kelsell, *et al.*, 1995) initially placed *RPGRIP on* chromosome 14. This was confirmed after completion of much of this work when the human chromosome 14 genomic sequence containing *RPGRIP* appeared in the genome databases. The gene was also localised to chromosome 14 on two human gene maps that had been constructed using analogous techniques: a radiation hybrid map and the chromosome 14 HAPPY map (Dear *et al.*, 1998). These were both produced by fragmenting the genome (by irradiation or shearing), and then estimating the order and distances between separate loci on the basis of the frequency of cosegregation: the closer two loci are in the genome the more frequently they will have been retained on the same fragment.

The location of the *RPGRIP* gene was determined by radiation hybrid mapping using the Genebridge 4 panel (Gyapay *et al.*, 1996). In this way it was placed on the framework of markers in-between D14S264 and D14S275 in chromosomal region 14q11. This places *RPGRIP* proximal to D14S275. The gene was placed 6.94 centiRays-$_{6000}$ (cR) from the closest marker, D14S264, with a LOD score of 2.03 relative to the next most likely placement. A LOD score of 2.03 means that this position is 107 times more likely to be the true position than the next most likely placement. The breakage distance of 6.94 cR translates into a physical distance of 1443.52 kb on the Genebridge 4 radiation hybrid map of chromosome 14, where 1 cR equates to 208 kb.

Identification of an STS (h14a1407) containing 320 bp of *RPGRIP* sequence located the gene on the HAPPY map, a framework of 1003 markers covering the whole of chromosome 14 (Dear *et al.*, 1998). In addition to the STSs generated by Dear *et al.* (1998), the chromosome 14 HAPPY map contains a number of markers which are present on the radiation hybrid and genetic maps to facilitate cross-referencing between maps. STS:h14a1407 is located on the chromosome 14 HAPPY map between markers D14S72 and D14S275. D14S72 and D14S264 are both proximal to D14S275, so the placement of *RPGRIP* in this interval is in agreement with the radiation hybrid data.

A further goal of these experiments was to evaluate the expression of *RPGRIP* in a variety of human and bovine tissues. Expression was detected in retina and testis only in human tissues but was seen to be more widespread in bovine (although this may reflect the rapidity with which bovine tissues were harvested after death and the consequent likelihood that RNAs are undegraded). Three northern blots containing human poly $(A)^+$ RNA were probed in an attempt to determine the size, number and tissue distribution of *RPGRIP* transcripts. Only one of these gave a strong positive result when two testis transcripts were detected (northern blot 1), one of 2.0 kb and a less abundant one of 3.1 kb. No signals were detected on this blot in spleen, thymus, prostate, ovary, small intestine, colon or peripheral leukocyte. A faint band of 6.2 kb was detected in the skeletal muscle lane of the second northern blot (northern blot 2) and two weak bands of 5 kb and 6.25 kb were observed in the retina lane of the third northern blot (northern blot 3). These were presumed to be non-specific, although addition of the twelve upstream GENSCAN-predicted exons to the 2894 bp *RPGRIP* sequence would produce a transcript of 4483 bp, and a 1.7 kb 5' untranslated region could raise this to 6.2 kb. Alternative splicing was evident in *RPGRIP* (transcripts lacking exons 5-7 and 12 were isolated) so it remains a possibility that some of the GENSCAN-predicted exons are part of the *RPGRIP* gene but are not transcribed in the tissues used for cDNA analysis. The retina that was used to prepare the retinal RNA on this

northern blot was obtained from a different source to the other tissues and the much stronger β-*actin* signal in retina compared with the other tissues may reflect this. All of the RNA samples used to prepare the third blot were purified from tissue samples that had been collected an average of 24 hours after death. This is sufficient time for RNase-mediated degradation of the RNA to occur. No retina-containing northern filters (or RNAs) are available commercially. With the ready availability of fresh bovine RNA, hybridising an *RPGRIP* probe to a bovine northern blot would be a valuable future experiment.

RPGRIP expression is readily detectable in human and bovine retina using more sensitive RT-PCR analyses. Expression was evident in all bovine tissues examined with the exception of pancreas, although this RNA sample also failed to produce a strong *G3PDH* control signal. From a range of human total RNA samples, *RPGRIP* was only detectable in retina and testis (two bands were amplified from retina RNA; one lacked exon 12, the other was intact). In both the human and the bovine RT-PCR experiments, two separate reactions were used to generate the *RPGRIP/bRPGRIP* and *G3PDH* products. These reactions utilised gene specific primers during the reverse transcriptase step (rather than oligo (dT) primers to generate a cDNA pool from which specific transcripts could be subsequently amplified). The difference between human and bovine tissues may reflect a genuine species difference or a difference in mRNA quality: human retina RNA could only be obtained 24 hours post-mortem compared with less than 30 minutes for bovine tissues. Alternatively, the bovine RT-PCR products may represent illegitimate transcription products (Chelly *et al.*, 1989) or a processed pseudogene. The latter possibility could be confirmed or excluded by sequencing the bovine RT-PCR products. Human *RPGRIP* transcripts were also amplified from human lymphoblastoid cell line derived total RNA. However, two rounds of PCR using nested primers were required in order to obtain visible bands, so increasing the risk of detecting illegitimate transcription products. Consequently, any conclusion that *RPGRIP* is normally expressed in lymphoblastoid cells must be treated with caution.

The large number of sequences deposited in public EST databases (for example 3,288,343 human ESTs in dbEST (http://www.ncbi.nlm.nih.gov/dbEST/) from several thousand separate libraries) provides a means of assessing the range of expression of a gene. *RPGRIP* ESTs came from: retina; foetal liver and spleen; testis; mixed foetal lung, testis and B-cell; and pooled germ cell tumour libraries (see Table 4.5). The absence of *RPGRIP* ESTs from adult libraries without a retina or testis component suggests that *RPGRIP* expression in adults may be limited to retina and testis. In addition, the low number of ESTs in the databases may indicate low expression of *RPGRIP*.

Bioinformatic analyses of the primary structure can potentially generate functional information about a novel protein. Structure predictions using peptide analysis programs suggested that RPGRIP is a soluble protein containing two coiled-coil domains (encompassing residues 14 to 78 and 150 to 200) and two C_2-domains (residues 416-532 and 611-716). The coiled-coil was first identified as the major structural component of a large group of fibrous proteins that include keratin, myosin and fibrinogen (Pauling and Corey, 1953; Crick 1953). Coiled-coils consist of a number of α-helices, generally between two and four, wound into a superhelix (Cohen and Parry, 1986). The constituent α-helices may be arranged in parallel or antiparallel. At the amino acid level, coiled-coils are made up of 7-residue repeats. If the positions in these heptad repeats are labelled *a* to *g*, then the side chains of the residues at positions *a* and *d* are generally non-polar/hydrophobic and form the helix interface (Figure 4.21A), and *b*, *c*, *e*, *f* and *g* are polar/hydrophilic and make up the solvent-exposed part of the coiled-coil (Lupas 1996; Figure 4.21B). Positions *e* and *g* are frequently occupied by charged residues (Beck *et al.*, 1997). The *a* and *d* residues form a non-polar band running along the surface and inclined around the axis of each α-helix (Cohen and Parry, 1986; Figure 4.21C). These non-polar side chains are packed in the core of the bundle of α-helices in a distinctive manner, known as 'knobs-into-holes', to stabilise the coiled-coil (Figure 4.21A). This consists of a side chain from one helix (knob) positioned in a space bordered by four side chains of the opposing helix (hole), immediately to the side of the equivalent side chain from the opposing helix (Lupas, 1996).

Coiled-coils have a numbers of different functions including the provision of spacers to separate the bacterial cell wall from the outer membrane, providing a scaffold for regulatory complexes (tropomyosin), forming a defensive surface for pathogens (Staphylococci M proteins), forming the so-called antiparallel 'thumbs' and 'arms' in several DNA and RNA processing enzymes that move, guide or grip the reaction substrates (e.g. DNA polymerase I), contributing to the cytoskeleton (intermediate filaments), and forming molecular levers (myosin) and stalks (kinesin, influenza haemagglutinin), and producing mechanically rigid structures such as blood clots (fibrin) or hair (keratin) (reviewed by Lupas, 1996). Coiled-coils have also been found in protein segments identified by deletion analysis as being responsible for oligomerization (Jasin *et al.*, 1984). This is only a few of many examples to illustrate the fact that coiled-coils are present in a diverse array of proteins. Further study will be required to determine the function of this organised tertiary structural feature in RPGRIP.

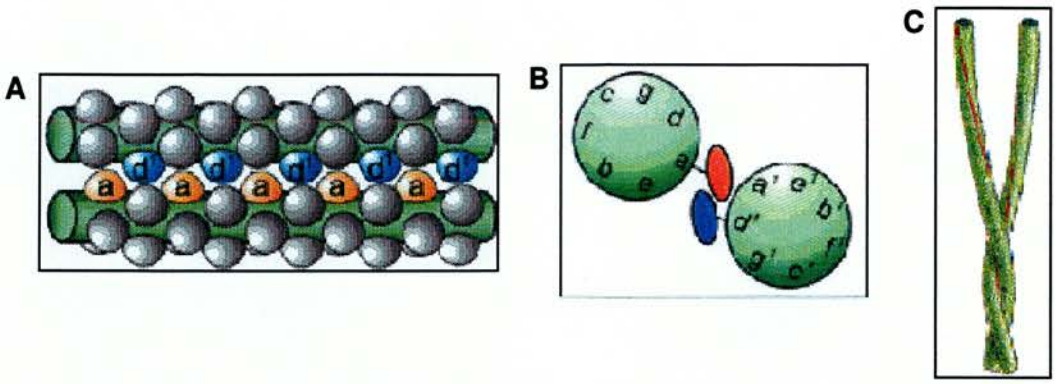


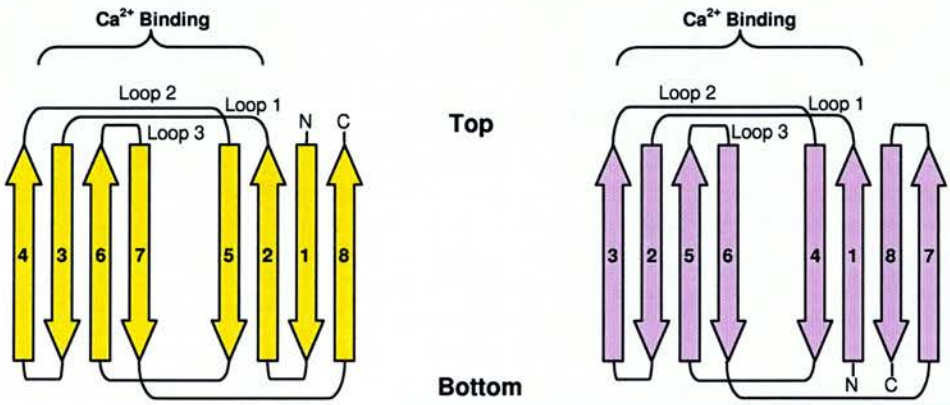
Figure 4.21: Models showing how two alpha helices interact to form a coiled-coil (Lumb and Kim, 1995). The primary sequence of the constituent helices consists of a repeat unit of seven residues, labelled *a* to *g* (**A** and **B**). The *a* and *d* residues tend to have non-polar/ hydrophobic side chains and form the interface between the two helices (**A**). The *b*, *c*, *e*, *f* and *g* residues generally have polar/hydrophilic side chains and make up the solvent-exposed part of the coiled-coil (**B**). The packing of the *a* and *d* non-polar side chains in the core of the α-helical bundle, known as 'knobs-into-holes', stabilises the coiled-coil. This consists of a side chain from one helix (knob) positioned in a space bordered by four side chains of the opposing helix (hole), immediately to the side of the equivalent side chain from the opposing helix (**A**). The *a* and *d* residues form a non-polar band running along the surface and inclined around the axis of each α-helix (represented by red stripes in **C**).

C_2-domains are independently folding calcium- and phospholipid-binding domains of approximately 130 amino acids. The majority of proteins with C_2-domains play roles in signal transduction, such as those involved in protein phosphorylation (e.g. protein kinase C (PKC) (Coussens *et al.*, 1986; Knopf *et al.*, 1986), ubiquitin ligation (e.g. Nedd4 (Plant *et al.*, 1997)), the activation of GTPases (e.g. Ras-GAP (Trahey et al., 1988)) and the generation of lipid second messengers (e.g. cytoplasmic phospholipase A_2 ($cPLA_2$) (Clark *et al.*, 1991), phospholipase Cs (PLCs; Rhee *et al.*, 1989), and phosphatidylinositol 3-kinases (Hiles *et al.*, 1992)). They are also involved in membrane trafficking, as with rabphilin-3 (Shirataki *et al.*, 1993), synaptotagmins (Perin *et al.*, 1990; Perin *et al.*, 1991), RIM (Wang *et al.*, 1997) and Munc13 (Brose *et al.*, 1995)) proteins.

The notion that C_2-domains may bind calcium ions followed from the observation that C_2-domain-containing PKC isoforms were regulated by Ca^{2+} whereas isoforms without C_2-domains were not (Nishizuka, 1988). For many C_2-domains, phospholipid binding is regulated by calcium and, for this reason, they are sometimes referred to as calcium-dependent lipid binding domains (reviewed by Rizo and Südhof, 1998).

Most of the information about C_2-domains has come from the study of synaptotagmin 1, $cPLA_2$, PKC, and inositol phospholipid-specific phospholipase C delta 1 subunit (PLCδ1). Synaptotagmin is one of at least twelve transmembrane proteins that contain two C_2-domains (the C_2A- and C_2B-domains) and plays a role in membrane

trafficking. It is found on synaptic vesicles and is required for the rapid Ca^{2+}-dependent element of the release of neurotransmitters (Geppert *et al.*, 1994; Brose *et al.*, 1992). A mechanism involving Ca^{2+} binding to both of the two synaptotagmin C_2-domains is thought to function as the chief calcium sensor in the exocytosis of synaptic vesicles. The C_2-domains in $cPLA_2$, PKC and PLCδ1 are thought to function by effecting the Ca^{2+}-dependent attachment of these proteins to phospholipid membranes (see Rizo and Südhof, 1998).



TOPOLOGY I
Synaptotagmin I,
PKCβ1

TOPOLOGY II
PLCδ1, $cPLA_2$

Figure 4.22: Schematic representations of the two defined C_2-domain topologies. The arrows represent b-strands and are numbered according to their appearance in the primary sequence. At the top of the C_2-domains the three calcium-binding loops are labelled loops 1-3.

In isolation, the synaptotagmin I C_2A-domain was found to display autonomous folding and Ca^{2+}-dependent phospholipid binding (Davletov and Südhof 1993). The domain was shown to bind all negatively charged phospholipids as well as syntaxin (Li *et al.*, 1995). The $cPLA_2$ (Nalefski *et al.*, 1994; Nalefski *et al.*, 1997), PKCβ (Shao *et al.*, 1996) and Nedd4 (Plant *et al.*, 1997) C_2-domains also bind phospholipids in a Ca^{2+}-dependent manner, although $cPLA_2$ binds to neutral phospholipids (Daveltov and Südhof 1993; Nalefski *et al.*, 1998) and PKCβ does not exhibit Ca^{2+}-dependent syntaxin 1 binding. These facts reflect the functional diversity among C_2-domains. Further to this, some Ca^{2+}-regulated C_2-domains simultaneously associate with other molecules independent of Ca^{2+} (e.g. synaptotagmin C_2B-domain interacts with AP-2 (Zhang *et al.*, 1994), inositol polyphosphates (Fukuda *et al.*, 1994), β-SNAP (Schiavo *et al.*, 1995) and Ca^{2+} channels (Sheng *et al.*, 1997)). Finally, the observation that a number of synaptotagmins seem to lack Ca^{2+} binding suggests that a number of C_2-domains may be Ca^{2+}-independent (see Rizo and Südhof, 1998).

The first studied structure of a C_2-domain was the C_2A-domain of synaptotagmin I (Sutton *et al.*, 1995). X-ray diffraction analysis showed this to consist of two closely apposed four-stranded β-sheets, with four loops at the bottom and three at the top connecting the eight strands (Figure 4.22). This conformation is referred to as topology I. The three-dimensional structure of the C_2-domain of PLCδ1 strongly resembles that of the synaptotagmin C_2A domain but the arrangement of the strands is very different (Essen *et al.*, 1996). This is a circular variation of topology I and is referred to as topology II (Figure 4.22). The PKC3 and $cPLA_2$ C_2-domain crystal structures were studied and found to be similar to those of the synaptotagmin I and PLCδ1 C_2-domains and found to display topologies I and II respectively (Perisic *et al.*, 1998). Binding sites for multiple (two or three) calcium ions are contained in the top loops of C_2-domains. These are formed for the most part by aspartate side chains, which serve as bidentate ligands for the Ca^{2+} ions. It appears that Ca^{2+}-dependent C_2-domains have evolved to cluster multiple Ca^{2+} ions within a small volume. Unsatisfied coordination sites on the Ca^{2+} ions are able to mediate interactions with molecular targets (see Rizo and Südhof, 1998). Comparison of the RPGRIP C_2-domain with those of other proteins suggests that it is likely to be a Ca^{2+}-binding domain (C. Ponting, personal communication).

Shao *et al.*, (1996) showed that calcium binding causes rotation of some side chains but induces no significant conformational change in synaptotagmin 1 C_2A-domain. This led to questions as to how Ca^{2+} regulates C_2-domain function. In the case of synaptotagmin 1 C_2A-domain, calcium binding leads to a significant change in the electrostatic potential of the protein (Shao *et al.*, 1996). It has been proposed that such changes may, in some cases, lead to target binding, including (for the synaptotagmin 1 C_2A-domain at least) phospholipid binding (Daveltov and Südhof 1993). However, work carried out on PKCβ has shown that lipid binding to C_2-domains is probably not entirely electrostatic (Edwards and Newton 1997), although what other factors influence this remain unknown.

Information about the potential function of a novel protein can also be gained by identifying homology with previously characterised sequences. BLAST searches were carried out in order to identify sequences with homology to RPGRIP. The sequence that was identified with the lowest e value (equivalent to the highest significance) was the uncharacterised human protein KIA1005. 'Weak similarity to myosin' had the next lowest e value and the subsequent sequences had much higher e values. This was the case for each of the three RPGRIP sequences that were used for database searching: (i) the 902 residue protein, (ii) the 1431 residue combination of the 902 amino acid RPGRIP and the 529 amino acids encoded by the GENSCAN-predicted exons and (iii) the 702 residue RPGRIP

sequence lacking the predicted coiled-coil domains. The e values are more significant for the top 10 matches with the 1431 residue RPGRIP sequence than with the other two query sequences. In the absence of experimental evidence as to which, if any, of the upstream GENSCAN-predicted exons are used or spliced, the *in silico* results obtained with the 1431 residue sequence should be treated with caution. However, the regions of homology with these sequences do not extend beyond the predicted first residue of the hypothetical 1259 amino acid RPGRIP protein identified as the likely orthologue of murine RPGRIP (described in section 4.6.3). Most of these homologous proteins are uncharacterised but three of them, hook1 protein, integrin homolog and intracellular transport protein USO1, are thought to play roles in intracellular transport, and two, paramyosin and "weak similarity to myosin", are potential motor proteins, suggesting the possibility of a similar role for RPGRIP.

Another aim of this chapter was to screen RPGRIP for mutations in RP patients. Bruford (1996) detected suggestive linkage to chromosome 14q11 in a subset of five out of eleven Sardinian ARRP families by typing and analysing 195 microsatellite markers. The *RPGRIP* gene was localised by radiation hybrid mapping to chromosome 14q11, close to the D14S275 marker Bruford identified as being linked to Sardinian ARRP. The RPGRIP gene was therefore screened for mutations in three of the five linked families but none were found. RPGRIP was amplified from RNA/cDNA and checked by genomic sequencing. The full genomic structure was obtained at a later stage (see section 4.3, page 121), after which exons 5, 6 and 7 were identified. Time constraints prevented these exons from being screened for mutations in the three ARRP families, and prevented the screening of RPGRIP in the other two linked families. This will need to be carried out in the future. If this fails to identify any mutations, and in the absence of other pedigrees linked to chromosomal region 14q11, a suitable approach would be to screen a large number of ARRP families in an attempt to detect *RPGRIP* mutations, which might be rare because ARRP is a very heterogeneous disorder with at least thirteen mapped or identified loci (see section 1.6, page 11).

In summary, a 2894 bp *RPGRIP* cDNA was isolated, which may be the full-length human transcript, and a smaller 1946 bp transcript, which lacks exons 5 to 7. The predicted products would yield proteins of 586 and 902 amino acids respectively, with molecular masses of 67 kD and 97 kD. The *RPGRIP* gene was localised to the long arm of chromosome 14 in band 14q11, between markers D14S264 and D14S275 and found to be expressed in retina and testis. The more widely expressed bovine homologue was detected in all except one of the tissues examined. Two potential coiled-coil and two C_2-domains were identified by *in silico* analysis of the RPGRIP protein sequence. These are likely to be important determinants of the function of RPGRIP.

CHAPTER FIVE

DISCUSSION

The principal aim of this project was to isolate proteins that interact with RPGR, the RP3 gene product. Using the yeast two-hybrid system, a previously uncharacterised sequence has been identified that encodes RPGR Interacting Protein (RPGRIP). Fifteen exons were shown to contribute to the *RPGRIP* gene, translating to a protein of 902 amino acids and 97 kDa. The bovine orthologue of *RPGRIP* (Roepman *et al.*, 2000) was reported at the same time as this work (Boylan and Wright, 2000) and the mouse orthologue was reported some time later (Hong *et al.*, 2000a). Comparison of the human RPGRIP protein with the bovine and mouse orthologues, indicates that additional upstream exons are likely to contain *RPGRIP* coding sequence. A GENSCAN analysis (Burge and Karlin, 1998) of human chromosome 14 genomic DNA identified 12 exons upstream of the presumed first exon of *RPGRIP* but these were not detected in human retina cDNA. *RPGRIP* is expressed in human testis (revealed by northern blot and RT-PCR) and retina (RT-PCR) but more widely in bovine tissues. The gene is located on chromosome 14q11 and preliminary screening *RPGRIP* in Sardinian ARRP patients, previously shown to be linked to this region, failed to identify any mutations. *In silico* analysis of RPGRIP revealed the presence of two potential coiled-coil domains, two potential calcium/phospholipid-binding C_2-domains and homology with characterised and uncharacterised proteins, including several with roles in intracellular trafficking.

The three research papers that reported the isolation of RPGRIP (Boylan and Wright, 2000; Hong *et al.*, 2000a; Roepman *et al.*, 2000) each presented the results of screening bovine or mouse retina expression libraries in the yeast two-hybrid system. Hong *et al.* (2000a) screened a mouse retina library using residues 39 to 460 of the mouse RPGR protein (mRPGR) as bait. They screened 1×10^6 library sequences and identified five mRPGRIP clones and one PDEδ clone, the latter replicating the result of Linari *et al.* (1999). The interaction between mRPGRIP and mRPGR was confirmed using a GST pull-down assay. Roepman *et al.* (2000) screened 1.85×10^6 independent bovine retina clones using residues 1 to 402 of human RPGR as bait. They identified six bovine *RPGRIP* clones and verified the interaction using a GST pull-down assay and co-immunoprecipitation of bRPGR and bRPGRIP from bovine retinal extracts using anti-RPGR and anti-RPGRIP antibodies.

By hybridising northern blots containing RNA purified from murine tissues to *mRPGRIP* probes, Hong *et al.* (2000a) detected a 10 kb doublet in retina and a 3 kb band in testis but no bands in brain, heart or liver, nor in retinal tissue from the *rd* mouse. Roepman

et al. (2000) probed bovine northern blots with *bRPGRIP* and observed a band of 0.8 kb in retina, brain, RPE, kidney, liver and spleen and retina-specific transcripts of 4.5 kb and 7.5 kb. RT-PCR experiments using bovine RNA samples revealed *RPGRIP* expression in heart, brain, spleen and retina but not liver, lung, skeletal muscle, placenta or kidney. Both groups also assessed *RPGRIP* expression in a variety of tissues by probing western blots with anti-RPGRIP antibodies. Hong *et al.* (2000a), using a polyclonal antibody that binds mRPGRIP residues 991-1331 (including the RPGR-binding domain), detected a 210 kDa polypeptide in the retina of wild-type and RPGR knockout mice but not in wild-type brain, heart, lung, muscle, testis, kidney or spleen. Roepman *et al.* (2000) used a polyclonal antibody against the RPGR-interacting domain of bRPGRIP (encompassing residues 842-1232) to probe western blots containing bovine proteins. They observed a 175 kDa retina-specific band; a 100 kDa band in retina, brain, skeletal muscle and lung; a 97 kDa band in spleen, kidney and lung and a 48 kDa band in retina, brain and skeletal muscle. This indicates either alternative splicing of *RPGRIP* in bovine tissues or differential protein degradation (possibly as a result of the extraction process). The results of expression studies described in the three reports are summarised in Table 5.1. The 2894 bp human *RPGRIP* sequence shown in figure 4.11 translates to a 902 amino acid protein with a predicted molecular weight of 97 kDa. This is likely to be homologous to either the 97 kDa isoform or the 100 kDa isoform observed by Roepman *et al.* (2000). The 1259 amino acid human protein identified in section 4.6.3 as being the probable human orthologue of bovine and mouse RPGRIP has a predicted molecular weight of 144 kDa. A protein encoded by a transcript containing exons 1 to 15 and GENSCAN-predicted exons -12 to -1 would have a molecular weight of 163 kDa. With significant post-translational modification, either the 144 kDa or 163 kDa hypothetical human proteins may be homologues of the 175 kDa bovine peptide observed by Roepman *et al.* (2000) or the 210 kDa mouse peptide observed by Hong *et al.* (2000a).

The antibody work of both groups extended to investigating the subcellular localisation of RPGRIP. Hong *et al.* (2000) carried out immunostaining of murine retinal sections and detected RPGRIP at the junction between photoreceptor inner and outer segments (staining was identical in wild-type and RPGR knockout mice). Immunostaining of mechanically disrupted photoreceptors revealed the localisation of mRPGRIP to the connecting cilium, with small amounts in the basal body. Similar results were obtained using cone-dense squirrel retinas, although the epitopes recognised by the anti-mRPGRIP antibody were not definitively shown to be present in this animal. Immunogold labelling showed mRPGRIP to be localised to the external side of the cilia microtubule doublets but not to the basal body. Hong *et al.* (2000b) had previously localised mRPGR to the photorec-

Table 5.1: *RPGRIP* expression was detected in a variety of mouse, human and bovine tissues by Hong *et al.* (2000a), Roepman *et al.* (2000) and Boylan and Wright (2000). Blank boxes indicate that *RPGRIP* expression was not investigated in these tissues. (Hum./Bov./Mur. north = sizes in kilobases of bands detected in human/bovine/mouse northern blots, Hum./Bov./Mur. RT = results of screening for *RPGRIP* transcripts by RT-PCR, Hum./Bov./Mur. west = sizes in kiloDaltons of proteins detected on western blots. S.I. = small intestine, P.B.L. = peripheral blood leukocyte. * A 10 kb doublet was observed)

	Hum. north	Hum. RT	Hum. west	Bov. north	Bov. RT	Bov. west	Mur. north	Mur. west
Adrenal		×			✓			
Brain		×		0.8		100, 48	×	×
Colon	×							
Heart		×			✓		×	×
Kidney		×		0.8	✓	97		×
Liver		×		0.8	✓		×	
Lung		×			✓	100, 97		×
Muscle					✓	100, 48		×
Ovary	×				✓			
Pancrea					×			
P.B.L.	×							
Placenta					×			
Prostate	×							
Retina		✓	100	7.5, 4.7, 0.8	✓	175, 100, 48	10*	210
RPE				0.8				
S.I.	×							
Spleen	×	×		0.8		97		×
Testis	2.0, 3.1	✓		✓	✓		3	×
Thymus	×							

eptor cilium using an antibody against the C-terminal 250 amino acids of mRPGR. By performing cell fractionation and extraction experiments on mouse photoreceptor inner and outer segments, Hong *et al.* (2000a) detected RPGRIP exclusively in a detergent-insoluble fraction that was enriched for the ciliary axoneme (the central fibrillar core of the cilium, consisting of 9 pairs of microtubules). In COS cells, recombinant mouse RPGRIP was observed to have a punctate and filamentous distribution and formed homo-dimeric or homo-oligomeric complexes. It did not co-localise with microtubules. These results led Hong *et al.* (2000a) to propose that RPGRIP is a structural component of the ciliary axoneme and that RPGR indirectly associates with the cilium via RPGRIP binding. Roepman *et al.* (2000) observed RPGRIP and RPGR (using the anti-RPGRIP antibody mentioned earlier and an antibody against residues 96-116 of RPGR) co-localising in the rod outer segments of bovine and human photoreceptors. No signals were detected in cone photoreceptors. They also reported a degree of staining at the junction between the rod inner and outer segments, suggestive of a ciliary localisation. Similar to Hong *et al.* (2000a), Roepman reported that a

transfected heterologous N-terminal RPGRIP polypeptide formed a visible and self-aggregated higher order complex. Roepman *et al.* (2000) also investigated the effects of XLRP mutations on hRPGRIP binding in the yeast two-hybrid system. The F130C, Q236X and 468:RNQIICX mutations abolished RPGRIP binding at 30°C and G275S decreased RPGRIP binding. A human RPGRIP fragment encompassing residues 579 to 762 (equivalent to residues 719 to 902 of the RPGRIP shown in figure 4.11) was identified as the minimal RPGR-binding domain by deletion mutational analysis. By radiation hybrid mapping, Roepman *et al.* (2000) mapped human *RPGRIP* to a locus defined by marker D14S932 in chromosomal region 14q11.

It is not clear what the function of RPGRIP is in the retina. However, from the data contained in the three reports (Boylan and Wright 2000; Hong *et al*, 2000a; Roepman *et al*, 2000) it is possible to speculate as to the roles of RPGRIP and RPGR and suggest experiments that would help to define them. Neither Hong *et al.* (2000a) nor Roepman *et al.* (2000) competed for antibody binding with peptide in order to exclude non-specific immunostaining. However, both detected RPGRIP in the photoreceptor cilium, which connects the inner and outer segments. The observation of *RPGRIP* transcripts in human testis and retina (sections 4.5.1.i, page 132, and 4.5.2, page 136) hinted at a ciliary localisation for RPGRIP as the spermatozoan flagellum is analogous to the photoreceptor cilium. The literature contains numerous examples of sperm and ciliary abnormalities accompanying retinitis pigmentosa (although some have questioned this phenomenon (Lee *et al.*, 1983; Szczesny, 1995)), including thinning of cilia (in XLRP, ADRP and ARRP; Szczesny, 1995), immotile cilia syndrome (often coupled with infertility due to impaired sperm velocity and motility; Ohga *et al.*, 1991; van Dorp *et al.*, 1992), abnormal structure and reduced sperm docosahexaenoic acid content (Connor *et al.*, 1997), abnormal nasal cilia (in ADRP, ARRP and Usher's syndrome; Arden and Fox, 1979) and abnormal ciliary axonemes (in XLRP; Hunter *et al.*, 1988) and at least one kindred with reported ciliary abnormalities is known to carry an *RPGR* mutation (Dry *et al.*, 1999).

The inner and outer segments are two morphologically and functionally distinct compartments of photoreceptor cells. Outer segment proteins are synthesised within the inner segments and the cilium must support their transport distally and prevent their retrograde diffusion (Spencer *et al.*, 1988). The functions of this organelle are yet to be comprehensively described but Wolfrum and Schmitt (2000) recently showed that after synthesis in the inner segment, rhodopsin, which accounts for up to 85% of outer segment protein (Molday, 1998), is transported through the connecting cilium prior to incorporation into the outer segment disc membranes. Rod outer segments are renewed continually and the

rate of addition of new membrane has been estimated at up to 3 μm^2/min (Besharse, 1986). RPE cells, which process the phagocytosed fragments of outer segments, are among the most phagocytically active cells in the human body (see section 1.3, page 6), reflecting this huge turnover of photoreceptor material. The connecting cilium is clearly a vital component of protein transport in photoreceptors. RPGR was previously localised to the connecting cilium and *RPGR* knockout mice displayed mis-localisation of blue and green cone opsins to the inner segment, perinuclear and synaptic regions, as well as showing reduced outer segment rhodopsin content, suggesting aberrant opsin trafficking (Hong *et al.*, 2000b). Moreover, the *in silico* analysis of RPGRIP presented in section 4.6.2 (page 141) revealed that RPGRIP is homologous to several proteins involved in intracellular transport (hook1 protein, integrin homolog and intracellular transport protein USO1). Therefore, the observations that RPGRIP interacts with RPGR and localises to the connecting cilium are compatible with the notion that it plays a role in intracellular trafficking between the inner and outer segments.

Vesicle-mediated intracellular protein transport is regulated by a complex molecular machinery involving an assortment of membrane-bound and soluble factors (see Guo *et al.*, 2000). Nascent proteins are directed to their final destination in transport vesicles after entering the endoplasmic reticulum lumen and traversing the Golgi complex for post-translational processing. Each step of vesicular transport consists of three stages: budding of the vesicle from donor membranes, targeting to new compartments and fusion of the vesicle with acceptor membranes (Rothman, 1994). The SNARE hypothesis (Rothman, 1994) proposes that at all stages of vesicular transport, vesicle and target membranes have one or more SNAP-receptor proteins (called v-SNAREs and t-SNAREs, respectively) that act as unique address markers. A vesicle covered in v-SNAREs will only fuse with a target membrane if it is displaying the cognate t-SNARE. The cytosolic '*N*-ethylmaleimide sensitive factor' (NSF) ATPase (Block *et al.*, 1988) binds to SNAREs through its interaction with soluble NSF attachment proteins (SNAPs, which interact with the membrane-bound SNARE proteins) to form a complex known as the '20S particle'. The ATPase activity of NSF subsequently causes the disassembly of the 20S particle and fusion of vesicle and target membranes (Sollner *et al.*, 1993). Most SNAREs are type II membrane proteins with short luminal or extracellular domains and large cytosolic domains. The cytosolic domains are proposed to mediate the SNAP-SNARE interaction and contain coiled-coils (Sollner and Rothman, 1996). Synaptotagmin, a well characterised C_2-domain containing protein, is an example of a v-SNARE (Südhof, 1995) and it is possible that RPGRIP, another C_2-domain/coiled-coil-containing protein is a v-SNARE also. Once fusion has occurred, v-SNAREs must be recycled back to the original donor compartment to preserve membrane

identity (Ferro-Novick and Jahn, 1994). It is not known how this is brought about but it is thought that v-SNAREs are inactive until they engage in an association with a specific Rab GTPase (Machamer, 1996). Transport of secretory vesicles is blocked by non-hydrolysable analogues of GTP (Rothman, 1994) and Rab proteins are known to be required for the assembly of SNARE complexes (Sollner and Rothman, 1996).

Rab proteins are Ras-related small GTPases that serve to regulate the movement of intracellular transport vesicles. Different Rab proteins regulate vesicular trafficking within distinct subcellular compartments (Novick and Brennwald, 1993) by switching between GDP- and GTP-bound forms (Zerial and Stenmark, 1993). A model has emerged describing the role Rab proteins play in regulating vesicular transport. In this, Rab-GDP is maintained in a complex with a guanine nucleotide dissociation inhibitor (GDI) protein in the cytosol. Donor membrane-associated Rab guanine nucleotide exchange factor (GEF) catalyses the exchange of GDP for GTP and subsequent dissociation of the GDI. This causes the unmasking of the Rab geranylgeranyl group, through which the GTPase becomes bound to the vesicle budding site on the membrane. The vesicle then buds from the donor membrane and moves through the cytoplasm to the target membrane. The vesicle-bound Rab-GTP is thought to mediate fusion with the target membrane by interacting with effector molecules present on the acceptor membrane. After GTP hydrolysis, which may be necessary for fusion or may occur after fusion, Rab-GDP re-establishes contact with GDI, dissociates from the target membrane and returns to the donor membrane (Novick and Zerial, 1997; Pfeffer, 1994; Nuoffer and Balch, 1994).

The accurate transport of newly-synthesised rhodopsin is essential for the maintenance of viable rod photoreceptors. Sorting of rhodopsin into specific post-Golgi membranes at the trans-Golgi network (TGN) involves the five C-terminal amino acids of the protein (Deretic *et al.*, 1998). Mutations affecting these five residues cause particularly severe forms of ADRP (Berson *et al.*, 1991; Sandberg *et al.*, 1995) and, in transgenic animals, defective delivery of rhodopsin to the rod outer segments (Li *et al.*, 1996; Sung *et al.*, 1994). Work has been done to characterise the factors regulating the intracellular transport of newly synthesised rhodopsin, particularly in frog photoreceptors (Deretic, 1998). In these cells, rhodopsin is synthesised on rough endoplasmic reticulum and then transported to the Golgi for processing. Post-Golgi vesicles bearing rhodopsin form at the trans-Golgi network, where they are separated from the vesicles that transport synaptophysin (an abundant synaptic vesicle protein) to the photoreceptor synapse (Deretic and Papermaster, 1995). Rabs 1, 6 and 8 have all been implicated in the trafficking of rhodopsin between the inner and outer segments. The locations within the retina of these proteins and the specific

effects of mutations in the genes encoding them help to interpret their functions. In transgenic dominant-negative *Drosophila* Rab1 mutants, the rough ER lumina became swollen, and the accumulation of immature rhodopsin and dissembly of Golgi bodies were noted (Satoh *et al.*, 1997). This suggests that Rab1 is required for the transport of newly synthesised proteins from the ER to the Golgi apparatus. Rab3 is found on rhodopsin-bearing post-Golgi vesicles, Golgi membranes and synaptic vesicles but appears to function in synaptic vesicle fusion rather than rhodopsin transport (Deretic and Papermaster, 1995). Expression of a GTPase-defective mammalian Rab6 protein led to a reduction in protein transport from the cis/medial to late Golgi compartments (Martinez *et al.*, 1994). In *Drosophila*, over-expression of a GTPase-defective Rab6 mutant led to reduced steady-state rhodopsin levels and prevented the maturation of this protein beyond an immature 40 kDa form (Shetty *et al.*, 1998). Post-translational modification of rhodopsin occurs in the cis- or medial-Golgi, so these experiments indicated that Rab6, in *Drosophila* at least, is involved in trafficking the nascent protein between the ER and Golgi. Rab6 has been detected on Golgi (Goud *et al.*, 1990), TGN (Antony *et al.*, 1992; Jones *et al.*, 1992) and post-Golgi membranes (Jasmin *et al.*, 1992) and in rod outer segments (Deretic and Papermaster, 1993). Depletion of Rab8 in developing hippocampal neuronal cells using anti-sense oligonucleotides led to a block in the morphological maturation of the cells and the inhibition of membrane traffic (Huber *et al.*, 1995). Rab8 is found on post-Golgi (Huber *et al.*, 1993a) and plasma membranes (Huber *et al.*, 1993b). In frog photoreceptors, Rab8 was localised to rod outer segments and to the base of the cilium where rhodopsin-bearing post-Golgi vesicles accumulate prior to their delivery to the outer segments (Deretic *et al.*, 1995). These localisations are shown schematically in Figure 5.1A.

The presence of an RCC1-homologous domain within the RPGR protein gave rise to speculation that it may act as a GEF for a small GTPase, since RCC1 performs this function on Ran. Choroideremia (see section 1.6.2.iii, page 19) is a similar disorder to RP and is caused by mutations in the *REP1* gene (Cremers *et al.*, 1990), which encodes subunit A of the Rab geranylgeranlytransferase enzyme (homologous to Rab GTP dissociation inhibitor). The resemblance of choroideremia to RP and the fact that the targets of the *REP1* gene product are the Rab GTPases, have led others to speculate that RPGR may act as the GEF for a Rab protein (Roepman *et al.*, 1996). It is possible that RPGR acts as the GEF for one of the Rab proteins found in the distal photoreceptor, such as Rab8, which is found in the connecting cilium (Deretic *et al.*, 1995). RPGRIP, which has been detected in the connecting cilium (Hong *et al.*, 2000a; Roepman *et al.*, 2000), and is possibly a structural component, may recruit RPGR to this compartment in order to play a role in the delivery of vesicles containing

rhodopsin, or other photoreceptor proteins, through the connecting cilium to the outer segment. GEFs appear to be required for vesicle budding in Rab GTPase-mediated vesicular transport (Rothman, 1994), i.e. the first of the three steps in movement of vesicles between compartments (budding, targeting, fusion). Because Rab6 and Rab8 have both been found in photoreceptor outer segments and connecting cilia (Deretic and Papermaster, 1993; Deretic *et al.*, 1995), it is conceivable that RPGR does possess GEF activity for one of these proteins (Figure 5.1).

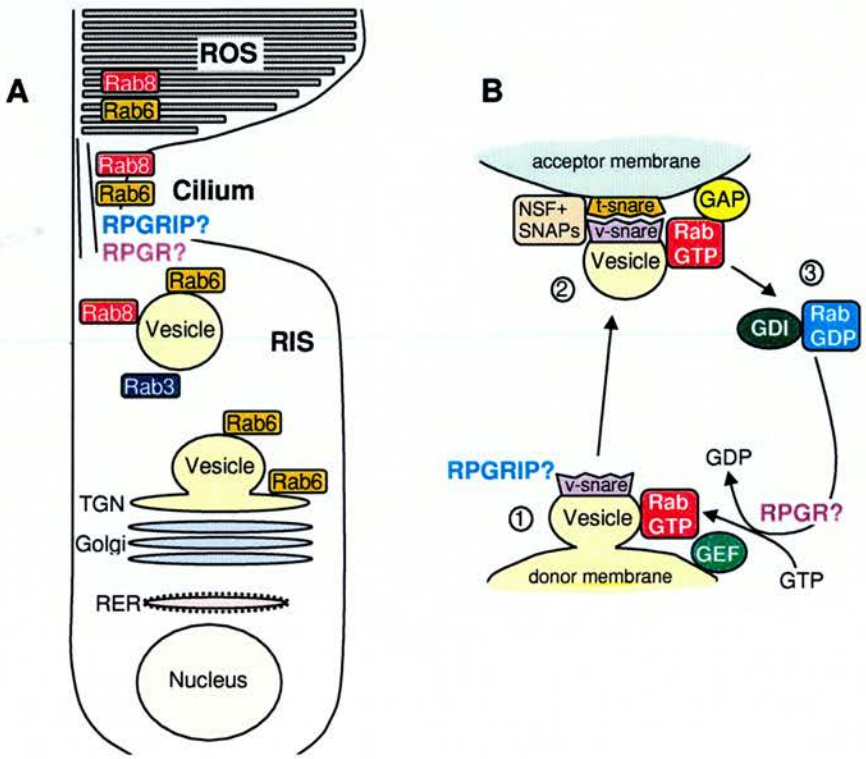


Figure 5.1: A. Photoreceptor compartments localisations to of Rab GTPases involved in post-Golgi trafficking of rhodopsin (Deretic and Papermaster, 1995) and potential localisation of RPGRIP (Roepman *et al.*, 2000; Hong *et al.*, 2000a) and RPGR (Hong *et al.*, 2000b). Rab8 and Rab6 are both found in outer segments, connecting cilia and on post-Golgi vesicles. ROS = rod outer segment, RIS = rod inner segment. **B.** Model of vesicular transport according to the SNARE hypothesis (Rothman, 1994; Deretic and Papermaster, 1995). (1) Protein vesicles bud from donor membranes. Inactive Rab-GDP binds to membrane-bound GEF (possibly RPGR) and Rab-GTP is formed. Rab-GTP binds to vesicle budding site, which also recruits v-SNARE membrane protein (possibly RPGRIP). Vesicle buds from the donor membrane and moves through cytoplasm to target membrane. (2) The v-SNARE recognises t-SNARE and this recognition initiates GTP hydrolysis, catalysed by acceptor membrane-associated Rab GTPase activating protein (GAP). Membrane fusion involves the 20S particle (NSF and SNAP proteins). (3) Inactive Rab-GDP is removed from the membrane by the Rab dissociation inhibitor (GDI) and recycled back to the donor compartment.

A potential RPGR/Rab interaction could be tested in the yeast two-hybrid system. However, the yeast two-hybrid system is not typically successful in showing GEF/GTPase interactions so direct testing of wild-type Rab and RPGR proteins is unlikely to give a positive result (De Toledo *et al.*, 2000). Two approaches may overcome this. The first would be to use dominant-negative Rab mutants (predominantly GDP-bound) in order to stabilise the Rab/GEF interaction. Other dominant-negative GTPases are known to bind to GEFs more readily, for example, the T24N Ran mutant is known to have an approximately 100-fold higher affinity for RCC1 than wild-type Ran (Lounsbury *et al.*, 1996). The second approach would be to use the yeast exchange assay, a modification of the yeast two-hybrid system, which was shown by de Toledo *et al.* (2000) to be useful for testing the specificity of newly identified GEF proteins. In this technique, three proteins are expressed in yeast two-hybrid host strains: a GTPase protein fused to the GAL4 DNA-binding domain, a GTPase effector protein fused to the GAL4 activation domain and the putative GEF protein. The specific Rab GTPase effector protein would be analogous to one of the Ran-binding proteins in the Ran/RCC1 system (Coutavas *et al.*, 1993) and its presence in the yeast nucleus makes the GTPase/GEF interaction much more readily detectable. The technique was used to identify a novel RhoA exchange factor using a Rho kinase effector and could be used to test candidate GTPases or screen yeast two-hybrid libraries. RPGR could also be tested for GTP/GDP exchange activity directly using GEF assays (Rosa and Barbacid, 1997). RPGR has been investigated in this way using a number of small GTPases but, to date, no GEF activity has been detected (Dr. F. Manson, personal communication).

Another potential future experiment would be a yeast two-hybrid library screen using RPGRIP as bait. The bovine retina library that isolated RPGRIP in the screen for RPGR-interacting proteins is a resource that could be used to characterise this novel protein. If RPGRIP is involved in intracellular trafficking then two-hybrid library screens may identify other proteins involved in this process. As Hong *et al.* (2000a) have suggested, RPGRIP may contribute to the structure of ciliary axonemes. If this is the case, then it is feasible for RPGR both to interact with a small GTPase involved in protein trafficking and to bind to RPGRIP. This is supported by the fact that Rab proteins are known to be intimately associated with cytoskeletal proteins and to regulate transport along microtubules (Somsel Rodman and Wandinger-Ness, 2000). Investigation of Rab8 and its yeast orthologue Sec4p, for example, have indicated that interactions with actin play a role in vesicle docking and fusion (Finger et al., 1998; Guo et al., 1999; Peränen et al., 1996). In addition, Deretic *et al.* (1995) speculated that Rab8 may mediate the attachment of rhodopsin vesicles to microfilament bundles prior to incorporation into rod outer segment membranes.

The C_2- and coiled-coil domains are likely to be important for the correct functioning of RPGRIP and should be investigated independently. These domains could be investigated in the yeast two-hybrid system in order to isolate interacting proteins that may participate in an RPGR/RPGRIP complex. The coiled-coil domains may mediate homo-oligomerisation, or interact with other retinal/photoreceptor proteins. The C_2-domain may bind Ca^{2+}, or bind phospholipids in a Ca^{2+}-dependent manner - activities which are not amenable to replication in the yeast two-hybrid system. However, the RPGRIP C_2-domains could be investigated for Ca^{2+}-dependent phospholipid binding using the method of Davletov and Südhof, (1993). In this experiment, recombinant GST-tagged RPGRIP C_2-domains would be attached to glutathione-agarose beads and incubated with radiolabelled liposomes (made from phosphatidylcholine mixed with phosphatidylserine, phosphatidylethanolamine or phosphatidylinositol, after labelling with 1,2-dipalmitoyl,L-3-phosphatidyl[N-*methyl*-^{3}H]choline). The degree of phospholipid binding in the presence of Ca^{2+} can be compared with binding in its absence by liquid scintillation counting of the collected agarose beads. Alternatively, the C_2-domains may mediate an interaction with other proteins and the yeast two-hybrid system could be used to investigate this.

The RPGR plaid domain is commonly mutated in XLRP patients and is likely to be present in retinal RPGR isoforms that also contain the RPGRIP-interacting/RCC1-homologous domain (Vervoort *et al.*, 2000). Screening of cDNA libraries in the yeast two-hybrid system for proteins interacting with the RPGR plaid domain would help to ascertain its role. Another use of the yeast two-hybrid system would be to define the minimal RPGRIP-interacting domain of RPGR in order to delineate its functional components.

The localisation of RPGR to the connecting cilium was achieved using an antibody against the C-terminal 250 amino acids of mRPGR (Hong *et al.*, 2000b). Vervoort *et al.* (2000) showed that RP3 mutations are predominantly located in the novel terminal exon ORF15 of *RPGR*. The protein encoded by a transcript containing this exon would not contain the epitopes recognised by the antibody used by Hong *et al.* (2000b). A useful future experiment would be to determine whether the subcellular localisation of the exon ORF15-encoded protein differs from that described by Hong *et al.* (2000b). The question of whether RPGRIP recruits RPGR to the connecting cilium, as Hong *et al.* (2000a) propose, could be investigated by immunostaining retinal tissue from an animal model of RP3 containing an *RPGR* mutation that abolishes binding to RPGRIP (e.g. G60V, H98Q, F130C, P235S, C250R). A fourteen-amino-acid peptide (equivalent to RPGRIP residues 796-809: DLDPQEQQGRRRFL) was synthesised and used to raise antibodies in two rabbits by Severn Biotech Ltd (Kidderminster, Worcester, UK) with a view to carrying out

immunohistochemical experiments. Unfortunately, there was insufficient time for antibody work to be carried out.

Human *RPGRIP* transcripts were detectable in only a very limited range of tissues (testis and retina), contrasting with apparently ubiquitous expression in bovine tissues. This may reflect the increased likelihood that human transcripts were degraded, since tissues were collected some time after death. The apparent differences in *RPGRIP* expression between bovine and human may also be due to RT-PCR amplification of illegitimate transcripts (Chelly *et al.*, 1989). If all transcripts are subject to RNase-degradation, the delayed extraction of human RNA would reduce the number of illegitimate transcripts in these tissues compared with bovine. Alternatively, bovine *RPGRIP* may be widely expressed and this is supported by the observation of Roepman *et al.* (2000) of a ubiquitous bovine transcript of 0.8 kb in addition to retina-specific transcripts. If fresh human samples could be obtained (especially difficult with retinas, which only rarely become available) then expression studies could be carried out with greater confidence. The absence of exons 5-7 from PCR-amplified *RPGRIP* transcripts, and of exon 12 from an RT-PCR-amplified retina transcript suggests alternative splicing patterns. Further cDNA analysis would help to characterise additional RPGRIP isoforms. Other retinal or tissue-specific exons or transcripts may throw light on potential RPGRIP functions. For example, the absence of specific *RPGRIP* exons in a transcript could remove one of the protein domains and affect its activity (e.g. exons 1, 2 and 4 encode the putative coiled-coils and exons 7, 9, 10 and 11 encode the putative C_2-domains). It would be a relatively straightforward task to check for the presence of each exon in mRNA transcripts from a range of tissues. None of the twelve GENSCAN-predicted exons were detected in human retina cDNA but comparison of human RPGR with the bovine and mouse orthologues strongly suggests that at least five of these (GENSCAN predicted exons -8 to -12) contribute to *RPGRIP*. An exhaustive, systematic cDNA analysis would determine which, if any, of these predicted exons are genuine, and in what tissues they are expressed.

The functions of RPGR in extra-ocular tissues are of interest since they may elucidate its functions in the retina. The ubiquitous nature of RPGR expression (Meindl *et al*, 1996) reflects that of PDEδ, the first RPGR-interacting protein to be identified (Linari *et al.*, 1999). Mutations in the PDEα and β subunits, but not the δ subunit, have been found in association with retinal degenerations, suggesting that PDEδ has a more general role outside the eye. The C-termini of RPGR isoforms encoded by exon 19-containing transcripts have a consensus isoprenylation site with the motif [CaaX] (where a is an aliphatic amino acid and X is leucine) preceded by a string of basic residues, as found in geranylgeranylated proteins

(Brown and Goldstein, 1993). This potentially enables RPGR to become anchored to membranes. It is conceivable therefore that PDEδ solubilises membrane-bound RPGR in the same way as it solubilises the PDE holoenzyme (Florio, Prusti, and Beavo, 1996). It is also possible that this is predominantly an extra-ocular process. Vervoort *et al.* (2000) have showed that the mutationally important part of RPGR, the plaid domain, is encoded by the terminal exon ORF15. Isoforms containing this domain will lack the isoprenylation motif and will not be membrane-bound in this tissue unless through interaction with other proteins. It may be possible to test the idea that PDEδ solubilises membrane-bound RPGR in a heterologous expression system. Yan *et al.* (1998) showed that, in transfected COS cells, mouse RPGR is isoprenylated and localised to the Golgi complex. It would be interesting to express the isoprenylated isoform of RPGR and note whether the addition of a PDEδ expression construct led to the solubilisation/removal of RPGR from the Golgi complex. The interaction with PDEδ may, however, be important for the normal function of RPGR in the retina and, in support of this, Linari *et al.* showed that a number of RP3 mutations disrupt RPGR/PDEδ binding. Moreover, PDEδ is known to bind to several small GTPases including Rab13 (Marzesco *et al.*, 1998), Rho6, Arl3 and Arl184 (Linari *et al.*, 1999). This again links RPGR (indirectly, through its interaction with PDEδ) with small GTPases and, potentially, with protein or vesicle transport.

Finally, an RPGRIP-deficient animal would be a valuable asset for investigating the subcellular localisation of RPGR, vesicular trafficking and photoreceptor function, etc. It would also be interesting to observe whether the animal displays any retinal disease. No mutations were found in *RPGRIP* during the preliminary screening of Sardinian ARRP patients (section 4.7, page 149) although screening of a large panel of RP patients may lead to the identification of RPGRIP mutations in association with retinal disease. A recent ARVO abstract details the screening of *RPGRIP* in unrelated patients with Leber's congenital amaurosis (McGee *et al.*, 2001). Four null alleles were identified in the 3 out of 57 (5.3%) patients: one was homozygous for the mutation Lys342(1-bp del A), one was homozygous for Gln893(1-bp ins T) and one was a compound heterozygote with the mutations Asp1176(1-bp del T) and Trp65Ter. The numbers refer to a 1259 amino acid protein that the authors describe as the full length RPGRIP and which is presumably identical to the 1259 residue protein described in section 4.6.3 (page 145) and shown in Figure 4.20 (page 147). These mutations all lead to truncated proteins (shown schematically in Figure 5.2) and this may be reflected in the severe phenotype. The mutations all remove at least one of the predicted internal domains from RPGRIP with the exception of Asp1176(1-bp del T). However, the region of RPGRIP that was shown to mediate the

interaction with RPGR is removed by this mutation, indicating that disruption of the interaction may be sufficient to cause retinal disease. It is possible that *RPGRIP* missense mutations that have a less dramatic effect on the protein might be found in recessive RP patients. Additional genes are mutated in LCA (see section 1.6.1.iv, page 15) and the figure of 5% of patients with *RPGRIP* mutations compares with 7-16% with *RPE65* mutations, approximately 6% with *GUCYD* mutations, approximately 3% with *CRX* mutations and approximately 10% with *AIPL1* mutations (all figures from RetNet, the Retinal Information Network http://www.sph.uth.tmc.edu/RetNet/).

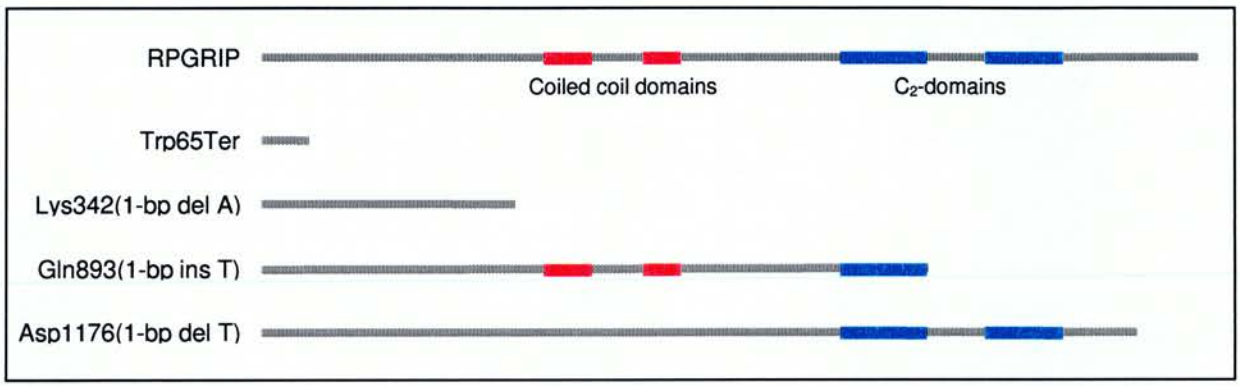


Figure 5.2: Schematic representation of truncated RPGRIP proteins resulting from null mutations found in LCA patients (McGee *et al.*, 2001). Grey lines indicate primary sequence of RPGRIP, red lines indicate the positions of the potential coiled coil domains, blue lines indicate the positions of the potential C_2-domains.

In summary, a new protein, called RPGRIP, has been identified as an interacting partner for the retinal protein RPGR. *RPGR* is mutated in up to 11% of all retinitis pigmentosa patients (Vervoort *et al.*, 2000) and the first comprehensive screening for *RPGRIP* mutations revealed genetic lesions in 5.3% of Leber's congenital amaurosis patients (McGee *et al.*, 2001). These results indicate that RPGR and the novel RPGRIP described here are fundamentally important for normal retinal function and that their further characterisation will aid the understanding of this tissue.

REFERENCES

Ahnelt, P.K. (1998). The photoreceptor mosaic. *Eye*, **12**, 531-40.

Alexandrov, K., Horiuchi, H., Steele-Mortimoer, O., Seabra, M.C. and Zerial, M. (1994). Rab escort protein-1 is a multifunctional protein that accompanies newly prenylated rab proteins to their target membranes. *EMBO J.*, **13**, 5262-5273.

Allikmets, R. and The International ABCR Screening Consortium (1999). Association of G1961E and D2177N variants in the ABCR gene with age-related macular degeneration. *Invest. Ophthalmol. Vis. Sci.*, **40**, S775.

Allikmets, R., Singh, N., Sun, H., Shroyer, N.F., Hutchinson, A., Chidambaram, A., Gerrard, B., Baird, L., Stauffer, D., Peiffer, A., Rattner, A., Smallwood, P., Li, Y., Anderson, K.L., Lewis, R.A., Nathans, J., Leppert, M., Dean, M. and Lupski, J.R. (1997a). A photoreceptor cell-specific ATP-binding transporter gene (ABCR) is mutated in recessive Stargardt macular dystrophy. *Nature Genet.*, **15**, 236-246.

Allikmets, R., Shroyer, N.F., Singh, N., Seddon, J.M., Lewis, R.A., Bernstein, P. S., Peiffer, A., Zabriskie, N.A., Hutchinson, A., Dean, M., Lupski, J.R. and Leppert, M. (1997b). Mutation of the Stargardt disease gene (ABCR) in age-related macular degeneration. *Science*, **277**, 1805-1807.

Altherr, M.R., Wasmuth, J.J., Seldin, M.F., Nadeau, J.H., Baehr, W. and Pittler, S.J. (1992). Chromosome mapping of the rod photoreceptor cGMP phosphodiesterase beta-subunit gene in mouse and human: tight linkage to the Huntington disease region (4p16.3). *Genomics*, **12**, 750-754.

Ammann, F., Klein, D. and Franceschetti, A. (1965). Genetic and epidemiological investigation of pigmentary degeneration of the retina and allied disorders in Switzerland. *J. Neurol. Sci.*, **2**, 183-196.

Ammerer, G. (1983). Expression of genes in yeast using the *ADC1* promoter. *Methods Enzymol.*, **101**, 192-201.

Anant, J.S., Ong, O.C., Xie, H.Y., Clarke, S., O'Brien, P.J. and Fung, B.K. (1992). In vivo differential prenylation of retinal cyclic GMP phosphodiesterase catalytic subunits. *J. Biol. Chem.*, **267**, 687-690.

Antonarakis, S.E. and McKusick, V.A. (2000). OMIM passes the 1,000-disease-gene mark. *Nature Genet.*, **25**, 11.

Antony, C., Cibert, C., Geraud, G., Santa Maria, A., Maro, B., Mayau, V. and Goud, B. (1992). The small GTP-binding protein rab6p is distributed from medial Golgi to the transGolgi network as determined by a confocal microscopic approach. *J. Cell Sci.***103**, 785-96.

Apte, S.S., Mattei, M.G. and Olsen, B.R. (1994). Cloning of the cDNA encoding human tissue inhibitor of metalloproteinases-3 (TIMP-3) and mapping of the TIMP3 gene to chromosome 22. *Genomics*, **19**, 86-90.

Arden, G.B. and Fox, B. (1979). Increased incidence of abnormal nasal cilia in patients with retinitis pigmentosa. *Nature*, **279**, 534-536.

Attwood, T., Croning, M., Flower, D., Lewis, A., Mabey, E., Scordis, P., Selley, J. and Wright, W. (2000). PRINTS-S: the database formerly known as PRINTS. *Nucleic Acids Res.*, **28**, 225-227.

Audic, S. and Claverie, J. (1998). Visualizing the competitive recognition of TATA-boxes in vertebrate promoters. *Trends Genet.*, **14**, 10-11.

Azizkhan, J.C., Jensen, D.E., Pierce, A.J. and Wade, M. (1993). Transcription from TATA-less promoters: dihydrofolate reductase as a model. *Crit. Rev. Eukaryotic Gene Expr.*, **3**, 229-254.

Banerjee, P., Kleyn, P.W., Knowles, J.A., Lewis, C.A., Ross, B.M., Parano, E., Kovats, S.G., Lee, J.J., Penchaszadeh, G.K., Ott, J., Jacobson, S.G. and Gilliam, T.C. (1998). TULP1 mutation in two extended Dominican kindreds with autosomal recessive retinitis pigmentosa. *Nature Genet.*, **18**, 177-179.

Bartel, P.L., Chien, C., Sternglanz, R. and Fields, S. (1993b). Elimination of false positives that arise in using the two-hybrid system. *BioTechniques*, **14**, 920-924.

Bartel, P.L., Chien, C., Sternglanz, R. and Fields, S. (1993a). Using the two hybrid system to detect protein-protein interaction. *In*: D.A. Hartley (ed.) *Cellular interactions in development: a practical approach*, pp.153-179, Oxford University Press, Oxford, UK.

Bascom, R.A., Garcia-Heras, J., Hsieh, C.L., Gerhard, D.S., Jones, C., Francke, U., Willard, H.F., Ledbetter, D.H. and McInnes, R.R. (1992). Localization of the photoreceptor gene ROM1 to human chromosome 11 and mouse chromosome 19: sublocalization to human 11q13 between PGA and PYGM. *Am. J. Hum. Genet.*, **51**, 1028-1035.

Bascom, R.A., Liu, L., Humphries, P., Fishman, G.A., Murray, J.C. and McInnes, R.R. (1993). Polymorphisms and rare sequence variants at the ROM1 locus. *Hum. Molec. Genet.*, **2**, 1975-1977.

Bateman, A., Birney, E., Durbin, R., Eddy, S., Howe, K. and Sonnhammer, E. (2000). The Pfam protein families database. *Nucleic Acids Res.*, **28**, 263-266.

Bech-Hansen, N.T., Naylor, M., Maybaum, T.A., Pearce, W.G., Koop, B., Fishman, G.A., Mets, M., Musarella, M.A. and Boycott, K.M. (1998). Loss-of-function mutations in a calcium-channel alpha-1 subunit gene in Xp11.23 cause incomplete X-linked congenital stationary night blindness. *Nature Genet.*, **19**, 264-67.

Beck, K., Gambee, J.E., Kamawal, A. and Bächinger, H.P. (1997). A single amino acid can switch the oligomerization state of the alpha-helical coiled-coil domain of cartilage matrix protein. *EMBO J.*, **16**, 3767-3777.

Bennett, J., Tanabe, T., Sun, D., Zeng, Y., Kjeldbye, H., Gouras, P. and Maguire, A.M. (1996). Photoreceptor cell rescue in retinal degeneration (rd) mice by in vivo gene therapy. *Nature Med.*, **2**, 649.

Béranger, F., Aresta, S., de Gunzburg, J. and Camonis, J. (1997). Getting more from the two-hybrid system: N-terminal fusion to LexA are efficient and sensitive baits for two-hybrid studies. *Nucleic Acids Res.*, **25**, 2035-2036.

Bergen, A.A.B., van den Born, L.I., Schuurman, E.J., Pinckers, A.J., van Ommen, G.J., Bleeker-Wagemakers, E.M. and Sandkuijl, L.A. (1995). Multipoint linkage analysis and homogeneity tests in 15 Dutch X-linked retinitis pigmentosa families. *Ophthal. Genet.*, **16**, 63-70.

Bernardinelli, L., Maida, A., Morinoni, A., Clayton, D., Romano, G., Montomoli, C., Fadda, D., Solinas, G., Castiglia, P., Cocco, P.L., Ghislandi, M., Berzuini, C., Pascutto, C., Nerini, M., Styles, B., Capocaccia, R., Lispi, L. and Mallardo, E. (1994). *Atlas of Cancer Mortality in Sardinia 1983-1987*, Progetto Finalizzato FATMA a cura della Direzione, Rome, Italy.

Berson, E.L., Rosner, B., Sandberg, M.A., Hayes, K.C., Nicholson, B.W., Weigel-Difranco, C. and Willett, W. (1993). A randomized trial of vitamin A and vitamin E supplementation for retinitis pigmentosa. *Arch. Ophthal.*, **111**, 761-772.

Berson, E. L., Rosner, B., Sandberg, M. A., Weigel DiFranco, C. & Dryja, T. P. (1991). Ocular findings in patients with autosomal dominant retinitis pigmentosa and rhodopsin, proline-347-leucine. *Am. J. Ophthalmol.* **111**, 614-623.

Besharse, J. C. (1986). *In*: Adler, R. and Farber, D. (eds.) The Retina: A Model for Cell Biological Studies, pp. 297-352, Academic, New York.

Bessant, D.A.R., Payne, A.M., Mitton, K.P., Wang, Q.-L., Swain, P.K., Plant, C., Bird, A.C., Zack, D.J., Swaroop, A. and Bhattacharya, S.S. (1999). A mutation in NRL is associated with autosomal dominant retinitis pigmentosa. *Nature Genet.*, **21**, 355-356.

Bhattacharya, S.S., Wright, A.F., Clayton, J.F., Price, W.H., Phillips, C.S., McKeown, C.M.E., Jay, M., Bird, A.C., Pearson, P.L., Southern, E.M. and Evans, H.J. (1984). Close genetic linkage between X-linked retinitis pigmentosa and a restriction fragment length polymorphism identified by recombinant DNA probe L1.28. *Nature*, **309**, 253-255.

Bird, A.C. and Heckenlively, J.R. (1988). X-linked recessive retinitis pigmentosa (X-linked pigmentary retinopathies). *In*: J.R. Heckenlively (ed.) *Retinitis Pigmentosa*, pp.162-187, J.B. Lippincott Company, Philadelphia, USA.

Bird, A.C. and Jay, B. (1994). Diagnosis in inherited retinal disorders. *In*: A.F. Wright and B. Jay (eds.) *Molecular genetics of inherited eye disorders*, pp.53-88, Harwood Academic Publishers, Chur, Switzerland.

Birnboim, H.C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.*, **7**, 1513-23.

Bischoff, F.R., Maier, G., Tilz, G. and Ponstingl, H. (1990). A 47-kDa human nuclear protein recognized by antikinetochore autoimmune sera is homologous with the protein encoded by RCC1, a gene implicated in onset of chromosome condensation. *Proc. Nat. Acad. Sci.* **87**, 8617-8621.

Bischoff, F.R. and Ponstingl, H. (1991). Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. *Nature*, **354**, 80-82.

Bischoff, F.R., Ponstingl, H., Coutavas, E. and D'Eustachio, P. (1996). Ran/TC4. *In*: M. Zerial and L. Huber (eds.) *Guidebook to the small GTPases*, pp. 457-460, Oxford Univserity Press/Sambrook and Tooze Publications, Oxford, UK.

Bitner-Glindzicz, M., Lindley, K.J., Rutland, P., Blaydon, D., Smith, V.V., Milla, P.J., Hussain, K., Furth-Lavi, J., Cosgrove, K.E., Shepherd, R.M., Barnes, P.D., O'Brien, R.E., Farndon, P.A., Sowden, J., Liu, X.-Z., Scanlan, M.J., Malcolm, S., Dunne, M.J., Aynsley-Green, A. and Glaser, B. (2000). A recessive contiguous gene deletion causing infantile hyperinsulinism, enteropathy and deafness identifies the Usher type 1C gene. *Nature Genet.*, **26**, 56-60.

Blacharski, P.A. (1988). Fundus Flavimaculatus. *In*: D.A. Newsome (ed.) *Retinal dystrophies and degenerations*, pp.135-159, Raven Press, New York, USA.

Block, M. R., Glick, B. S., Wilcox, C. A., Wieland, F. T. and Rothman, J. E. (1988) Purification of an *N*-ethylmaleimide-sensitive protein catalyzing vesicular transport. *Proc. Natl. Acad. Sci. USA* **85**, 7852-7856.

Boggon, T.J., Shan, W.S., Santagata, S., Myers, S.C. and Shapiro, L. (1999). Implication of tubby proteins as transcription factors by structure-based functional analysis. *Science*, **286**, 2119-25.

Bolz, H., von Brederlow, B., Ramirez, A., Bryda, E.C., Kutsche, K., Nothwang, H.G., Seeliger, M., Cabrera, M.C.-S., Vila, M.C., Molina, O.P., Gal, A. and Kubisch, C. (2001). Mutation of CDH23, encoding a new member of the cadherin gene family, causes Usher syndrome type 1D. *Nature Genet.*, **27**, 108-112.

Bonneau, D., Kaplan, J., Girard, G. and Dufier, J.L. (1992). Autosomal inheritance of "senile" retinitis pigmentosa: a report of a family with consanguinity. *Clin. Genet.*, **42**, 199-200.

Bonneau, D., Raymond, F., Kremer, C., Klossek, J.M., Kaplan, J. and Patte, F. (1993). Usher syndrome type I associated with bronchiectasis and immotile nasal cilia in two brothers. *J. Med. Genet.* **30**, 253-4.

Bork, J.M., Peters, L.M., Riazuddin, S., Bernstein, S.L., Ahmed, Z.M., Ness, S.L., Polomeno, R., Ramesh, A., Schloss, M., Srisailpathy, C.R.S., Wayne, S., Bellman, S., Desmukh, D., Ahmed, Z., Khan, S.N., Kaloustian, V.M., Li, X.C., Lalwani, A., Riazuddin, S., Bitner-Glindzicz, M., Nance, W.E., Liu, X.Z., Wistow, G., Smith, R.J., Griffith, A.J., Wilcox, E.R., Friedman, T.B. and Morell, R.J. (2001). Usher syndrome 1D and nonsyndromic autosomal recessive deafness DFNB12 are caused by allelic mutations of the novel cadherin-like gene CDH23. *Am. J. Hum. Genet.*, **68**, 26-37.

Boughman, J.A. and Fishman, G.A. (1983). A genetic analysis of retinitis pigmentosa. *Br. J. Ophthalmol.*, **67**, 449-454.

Boughman, J.A., Conneally, P.M. and Nance, W.E. (1980). Population genetic studies of retinitis pigmentosa. *Am. J. Hum. Genet.*, **32**, 223-235.

Boughman, J.A., Vernon, M. and Shaver, K.A. (1983). Usher syndrome: definition and estimate of prevalence from two high-risk populations. *J. Chronic Dis.*, **36**, 595-603.

Bowes, C., Li, T., Danciger, M., Baxter, L.C., Applebury, M.L. and Farber, D.B. (1990). Retinal degeneration in the rd mouse is caused by a defect in the beta subunit of rod cGMP-phosphodiesterase. *Nature*, **347**, 677-680.

Boylan, J.P., Wright, A.F. (2000). Identification of a novel protein interacting with RPGR. *Hum. Mol. Genet.* **9**, 2085-93.

Breuer, D.K., Affer, M., Andreasson, S., Birch, D.G., Fishman, G.A., Heckenlively, J.R., Hiryanna, S., Hoffman, D.R., Jacobson, S.G., Mears, A.J., Musarella, M.A., Redolfi, B., Sieving, P.A., Wright, A.F., Yashar, B.M., Zucchi, I and Swaroop, A. (2001). X-linked retinitis pigmentosa: current status. Epub ahead of print.

Breuer, D.K., Musarella, M. and Swaroop, A. (2000). Verification and fine mapping of the X-linked retinitis pigmentosa locus RP6. *Invest. Ophthalmol. Vis. Sci.*, **41**, S191.

Brose, N., Petrenko, A.G., Südhof, T.C. and Jahn, R. (1992). Synaptotagmin: a calcium sensor on the synaptic vesicle surface. *Science*, **256**, 1021-1025.

Brose, N., Hofmann, K., Hata, Y. and Sudhof, T.C. (1995) .Mammalian homologues of Caenorhabditis elegans unc-13 gene define novel family of C2-domain proteins. *J. Biol. Chem.***270**, 25273-80.

Brown, J., Dry, K.L., Edgar, A.J., Pryde, F.E., Hardwick, L.J., Aldred, M.A., Lester, D.H., Boyle, S., Kaplan, J., Dufier, J.L., Ho, M.F., Monaco, A.M., Musarella, M.A. and Wright, A.F. (1996). Analysis of three deletion breakpoints in Xp21.1 and the further localization of RP3. *Genomics*, **37**, 200-10.

Brown, M.S. and Goldstein, J.L. (1993) Mad bet for rab. *Nature*, **366**, 14-15.

Bruford, E. (1996). A genetic analysis of autosomal recessive forms of retinitis pigmentosa. University of Edinburgh PhD thesis.

Bruford, E.A., Mansfield, D.C., Teague, P.W., Barber, A., Fossarello, M. and Wright, A.F. (1994). Genetic linkage studies in autosomal recessive retinitis pigmentosa. *Am. J. Hum. Genet.*, **55**, A181.

Bundey, S. and Crews, S.J. (1984a). A study of retinitis pigmentosa in the city of Birmingham. I Prevalence. *J. Med. Genet.*, **21**, 417-420.

Bundey, S. and Crews, S.J. (1984b). A study of retinitis pigmentosa in the city of Birmingham. II Clinical and genetic heterogeneity. *J. Med. Genet.*, **21**, 421-428.

Bunker, C.H., Berson, E.L., Bromley, W.C., Hayes, R.P. and Roderick, T.H. (1984). Prevalence of retinitis pigmentosa in Maine. *Am. J. Ophthalmol.*, **97**, 357-365.

Buraczynska, M., Wu, W., Fujita, R., Buraczynska, K., Phelps, E., Andreasson, S., Bennett, J., Birch , D.G., Fishman, G.A., Hoffman, D.R., Inana, G., Jacobson, S.G., Musarella, M.A., Sieving, P.A. and Swaroop, A. (1997). Spectrum of mutations in the RPGR gene that are identified in 20% of families with X-linked retinitis pigmentosa. *Am. J. Hum. Genet.*, **61**, 1287-92.

Burge, C. and Karlin, S. (1998). Finding the genes in genomic DNA. *Curr. Opin. Struct. Biol.*, **8**, 346-354.

Burstedt M.S.I., Sandgren, O., Holmgren, G. and Forsman-Semb, K. (1999). Bothnia dystrophy caused by mutations in the cellular retinaldehyde-binding protein gene (RLBP1) on chromosome 15q26. *Invest. Ophthalmol. Vis. Sci.*, **40**, 995-1000.

Calabrese, G., Sallese, M., Stornaiuolo, A., Stuppia, L., Palka, G. and De Blasi, A. (1994). Chromosome mapping of the human arrestin (SAG), beta-arrestin 2 (ARRB2), and beta-adrenergic receptor kinase 2 (ADRBK2) genes. *Genomics*, **23**, 286-288.

Cavalli-Sforza, L.L. and Piazza, A. (1993). Human genomic diversity in Europe: a summary of recent research and prospects for the future. *Eur. J. Hum. Genet.*, **1**, 3-18.

Chelly, J., Concordet, J.P., Kaplan, J.C. and Kahn A. (1989). Illegitimate transcription: transcription of any gene in any cell type. *Proc. Nat. Acad. Sci.*, **86**, 2617-21.

Chen, J.D., Halliday, F., Keith, G., Sheffield, L., Dickinson, P., Gray, R., Constable, I. and Denton, M. (1989). Linkage heterogeneity between X-linked retinitis pigmentosa and a map of 10 RFLP loci. *Am. J. Hum. Genet.*, **45**, 401-411.

Chen, S., Wang, Q.-L., Nie, Z., Sun, H., Lennon, G., Copeland, N.G., Gilbert, D.T., Jenkins, N.A. and Zack, D.J. (1997). Crx, a novel Otx-like paired-homeodomain protein, binds to and transactivates photoreceptor cell-specific genes. *Neuron*, **19**, 1017-1030.

Chien, C., Bartel, P., Sternglanz, R. and Fields, S. (1991). The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Nat. Acad. Sci.*, **88**, 9578-9582.

Chomczynski, P. and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. *Analytical Biochem.*, **162**, 156-159.

Chong, N.H., Alexander, R.A., Gin, T., Bird, A.C. and Luthert, P.J. (2000). TIMP-3, collagen, and elastin immunohistochemistry and histopathology of Sorsby's fundus dystrophy. *Invest. Ophthalmol. Vis. Sci.*, **41**, 898-902.

Clark, J.D., Lin, L.-L., Kriz, R.W., Ramesha, C.S., Sultzman, L.A., Lin, A.Y., Milona, N. and Knopf, J.L. (1991). A novel arachidonic acid-selective cytosolic PLA2 contains a Ca(2+)-dependent translocation domain with homology to PKC and GAP. *Cell*, **65**, 1043-1051.

Cohen, C. and Parry, D.A.D. (1986). α-helical coiled coils – a widespread motif in proteins. *Trends Biochem. Sci.*, **11**, 245-248.

Connor, W.E., Weleber, R.G., Defrancesco, C., Lin, D.S. and Wolf, D.P. (1997). Sperm abnormailities in retinitis pigmentosa. *Invest. Ophthalmol. Vis. Sci.* **38**, 2619-28.

Cook, T.A., Ghomashchi, F., Gelb, M.H., Florio, S.K. and Beavo, J.A. (2000). The delta subunit of type 6 phosphodiesterase reduces light-induced cGMP hydrolysis in rod outer segments. *J. Biol. Chem.*, **276**, 5248-5255.

Cooper, D.N. and Krawczak, M. (1993). *Human Gene Mutations*, Bios, Oxford, UK.

Cordenonsi, M., D'Atri, F., Hammar, E., Parry, D.A., Kendrick-Jones, J., Shore, D. and Citi, S. (1999). Cingulin contains globular and coiled-coil domains and interacts with ZO-1, ZO-2, ZO-3, and myosin. *J. Cell Biol.*, **147**, 1569-82.

Corpet, F., Gouzy, J. and Kahn, D. (1999). Recent improvements of the ProDom database of protein domain families. *Nucleic Acids Res.*, **27**, 263-267.

Coussens, L., Parker, P.J., Rhee, L., Yang-Feng, T.L., Chen, E., Waterfield, M.D., Francke, U. and Ullrich, A. (1986). Multiple distinct forms of bovine and human protein kinase C suggest diversity in cellular signaling pathways. *Science*, **233**, 859-866.

Coutavas, E., Ren, M., Oppenheim, J., D'Eustachio, P. and Rush, M. (1993). Characterisation of proteins that interact with the cell cycle regulatory protein Ran/TC4. *Nature* **366,** 585-7.

Cremers, F.P.M. and Rogers, H.H. (1995). Choroideremia. *In*: C.R. Scriver, A.L.Beaudet, W.S. Sly, D. Valle (eds.) *The Metabolic and Molecular Bases of Inherited Disease*, 7th Ed., pp.4311-4323, McGraw-Hill, New York, USA.

Cremers, F.P.M., van de Pol, D.J.R., van Kerkhoff, L.P.M., Wieringa, B. and Ropers, H.-H. (1990). Cloning of a gene that is rearranged in patients with choroideraemia. *Nature*, **347**, 674-677.

Crick, F.H.C. (1953). The packing of alpha helices.Simple coiled-coils. *Acta cystallog.* **6**, 689-697.

David, G., Abbas, N., Stevanin, G., Durr, A., Yvert, G., Cancel, G., Weber, C., Imbert, G., Saudou, F., Antoniou, E., Drabkin, H., Gemmill, R., Giunti, P., Benomar, A., Wood, N., Ruberg, M., Agid, Y., Mandel, J.L. and Brice, A. (1997). Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. *Nat. Genet.* **17**, 65-70.

Davletov, B.A. and Südhof, T. C. (1993). A single C_2 domain from synaptotagmin I is sufficient for high affinity Ca^{2+}/phospholipid binding. *J. Biol. Chem.* **268**, 26836-26390.

Davletov, B.A. and Südhof, T.C. (1993). A single C_2 domain from synaptotagmin I is sufficient for high affinity Ca^{2+}/ phospholipid binding. *J. Biol. Chem.*, **268**, 26386-26390.

D'Cruz, P.M., Yasumura, D., Weir, J., Matthes, M.T., Abderrahim, H., LaVail, M.M. and Vollrath, D. (2000). Mutation of the receptor tyrosine kinase gene MERTK in the retinal dystrophic RCS rat. *Hum. Mol. Genet.*, **9,** 645-51.

De Saint-Basile, G., Bohler, M.C., Fischer, A., Cartron, J., Dufier, J.L., Griscelli, C. and Orkin, S.H. (1988). Xp21 DNA microdeletion in a patient with chronic granulomatous disease, retinitis pigmentosa and McLeod phenotype. *Hum. Genet.*, **80**, 85-89.

De Toledo, M., Colombo, K., Nagase, T., Ohara, O., Fort, P. and Blangy, A. (2000). The yeast exchange assay, a new complementary method to screen for Dbl-like protein specificity: identification of a novel RhoA exchange factor. *FEBS letters* **480**, 287-292.

Dear, P., Bankier, A. and Piper, M. (1998). A high-resolution metric HAPPY map of human chromosome 14. *Genomics*, **48**, 232-241.

den Hollander, A.I., ten Brink, J.B., de Kok, Y.J., van Soest, S., van den Born, L.I., van Driel, M.A., van de Pol, D.J., Payne, A.M., Bhattacharya, S.S., Kellner, U., Hoyng, C.B., Westerveld, A., Brunner, H., Bleeker-Wagemakers, E.M., Deutman, A.F., Heckenlively, J.R., Cremers, F.P. and Bergen, A.A. (1999). Mutations in a human homologue of Drosophila crumbs cause retinitis pigmentosa (RP12). *Nature Genet.*, **23**, 1061-4036.

Deretic, D. (1998). Post-Golgi trafficking of rhodopsin in retinal phtotoreceptors. *Eye* **12**, 526-30.

Deretic, D. and Papermaster, D.S. (1993). Rab6 is associated with a compartments that transports rhodopsin from the trans Golgi to the site of rod outer segment disk formation in frog retinal photoreceptors. *J. Cell Sci.* **106**, 803-813.

Deretic, D. and Papermaster, D.S. (1995). The role of small G-proteins in the transport of newly synthesized rhodopsin. *Prog. Ret. Eye Res.* **14**, 249-65.

Deretic, D., Huber, L.A., Ransom, N., Mancini, M., Simons, K. and Papermaster, D.S. (1995) rab8 in retinal photoreceptors may participate in rhodopsin transport and in rod outer segment disk morphogenesis. *J.Cell Sci.* **108**, 215-224.

Deretic, D., Schmerl, S., Hargrave, P.A., Arendt, A. and McDowell, J.H. (1998). Regulation of sorting and post-Golgi trafficking of rhodopsin by its C-terminal sequence QVS(A)PA. *PNAS* **95**, 10620-10625.

Desai, A. and Hyman, A. (1999). Microtubule cytoskeleton: no longer an also Ran. *Curr. Biol.*, **9**, R704-R707.

Deutmann, A.F. (1977). Rod-cone dystrophy: primary, hereditary, pigmentary retinopathy, RP. *In*: A.E. Krill (ed.) *Hereditary Retinal and Choroidal Diseases*, pp.479-576, Harper and Row Publishers Inc., Hagerston, Maryland, USA.

Dodge, G.R., Kovalszky, I., McBride, O.W., Yi, H.F., Chu, M., Saitta, B., Stokes, D.G. and Iozzo, R.V. (1991). Human clathrin heavy chain (CLTC): partial molecular cloning, expression, and mapping of the gene to human chromosome 17q11-qter. *Genomics*, **11**, 174-178.

Drivas, G.T., Shih, A., Coutavas, E., Rush, M.G. and D'Eustachio, P. (1990). Characterization of four novel ras-like genes expressed in a human teratocarcinoma cell line. *Mol. Cell. Biol.*, **10**, 1793-1798.

Dry, K.L., Aldred, M.A., Edgar, A.J., Brown, J., Manson, F.D., Ho, M.F., Prosser, J., Hardwick, L.J., Lennon, A.A., Thomson, K., van Keuren, M., Kurnit, D.M., Bird, A.C., Jay, M., Monaco, A.P. and Wright, A.F. (1995). Identification of a novel gene, ETX1 from Xp21.1, a candidate gene for X-linked retintis pigmentosa (RP3). *Hum. Mol. Genet.*, **4**, 2347-53.

Dry, K.L., Manson, F.D., Lennon, A., Bergen, A.A., vanDorp, D.B. and Wright, A.F. (1999). Identification of a 5' Splice Site Mutation in the RPGR Gene in a Family With X-Linked Retinitis Pigmentosa (RP3). *Hum. Mutat.* **13**, 141-145.

Dryja, T.P. (1992). Doyne Lecture: Rhodopsin and autosomal dominant retinitis pigmentosa. *Eye*, **6**, 1-10.

Dryja, T.P. and Li, T. (1995). Molecular genetics of retinitis pigmentosa. *Hum. Mol. Genet.*, **4**, 1739-1743.

Dryja, T.P., Finn, J.T., Peng, Y.-W., McGee, T.L., Berson, E.L. and Yau, K.-W. (1995). Mutations in the gene encoding the alpha subunit of the rod cGMP-gated channel in autosomal recessive retinitis pigmentosa. *Proc. Nat. Acad. Sci.*, **92**, 10177-10181.

Dryja, T.P., Hahn, L.B., Kajiwara, K. and Berson, E.L. (1997). Dominant and digenic mutations in the peripherin/RDS and ROM1 genes in retinitis pigmentosa. *Invest. Ophthalmol. Vis. Sci.*, **38**, 1972-1982.

Dryja, T.P., Hahn, L.B., Reboul, T. and Arnaud, B. (1996). Missense mutation in the gene encoding the alpha subunit of rod transducin in the Nougaret form of congenital stationary night blindness. *Nature Genet.*, **13**, 358-365.

Dryja, T.P., McEvoy, J.A., McGee, T.L. and Berson, E.L. (2000). Novel rhodopsin mutations Gly114Val and Gln184Pro in dominant retinitis pigmentosa. *Invest. Ophthalmol. Vis. Sci.*, **41**, 3124-7.

Dryja, T.P., Rucinski, D.E., Chen, S.H. and Berson, E.L. (1999). Frequency of mutations in the gene encoding the alpha subunit of rod cGMP-phosphodiesterase in autosomal recessive retinitis pigmentosa. *Invest. Ophthalmol. Vis. Sci.*, **40**, 1859-65.

Durfee, T., Becherer, K., Chen, P, Yeh, S., Yang, Y., Kilburn, A.E., Lee, W. and Elledge, S.J. (1993). The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. *Genes Dev.*, **7**, 555-569.

Edwards, A.S. and Newton, A.C. (1997). Regulation of protein kinase C betaII by its C_2 domain. *Biochemistry*, **36**, 15615-15623.

El-Amraoui, A., Sahly, I., Picaud, S., Sahel, J., Abitbol, M. and Petit, C. (1996). Human Usher 1B/mouse *shaker-1*: the retinal phenotype discrepancy explained by the presence/ absence of myosin VIIA in the photoreceptor cells. *Hum. Mol. Genet.*, **5**, 1171-1178.

Ellis, D.S. and Heckenlively, J.R. (1988). Retinitis punctata albescens. *In*: J.R. Heckenlively (ed.) *Retinitis Pigmentosa*, pp.155-161, J.B. Lippincott Company, Philadelphia, USA.

Essen, L.-O., Perisic, O., Cheung, R., Katan, M. and Williams, R.L. (1996). Crystal structure of a mammalian phosphoinositide-specific phospholipase C delta. *Nature*, **380**, 595-602.

Eudy, J.D., Weston, M.D., Yao, S., Hoover, D.M., Rehm, H.L., Ma-Edmonds, M., Yan, D., Ahmad, I., Cheng, J.J., Ayuso, C., Cremers, C., Davenport, S., Moller, C., Talmadge, C.B., Beisel, K.W., Tamayo, M., Morton, C.C., Swaroop, A., Kimberling, W.J. and Sumegi, J. (1998). Mutation of a gene encoding a protein with extracellular matrix motifs in Usher syndrome type IIa. *Science*, **280**, 1753-1757.

Evans, J. (1995). Causes of blindness and partial sight in England and Wales 1990-1991. *Studies on Medical and Population Subjects* **No. 57**, pp.1-28, OPCS, HMSO, London.

Feinberg, A.P. and Vogelstein, B.(1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.*, **132**, 6-13.

Ferro-Novick, S. and Jahn, R.Vesicle fusion from yeast to man. *Nature* **370**, 191-3.

Fields, S. and Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Nature*, **340**, 245-246.

Fields, S. and Sternglanz, R. (1994). The two-hybrid system: an assay for protein-protein interactions. *Trends Genet.*, **10**, 286-292.

Finger, F. P., Hughes, T. E. and Novick, P. (1998). Sec3p is a spatial landmark for polarized secretion in budding yeast. *Cell* **92**, 559-571.

Fishman, G.A. (1978). Retinitis Pigmentosa genetic percentages. *Arch. Ophthalmol.* **96**, 822-826.

Fishman, G.A., Weinberg, A.B. and McMahon, T.T.(1986). X-linked recessive retinitis pigmentosa. Clinical characteristics of carriers. *Arch. Opthalmol.*, **104**, 1329-1335.

Florio, S.K., Prusti, R.K. and Beavo, J.A. (1996). Solubilisation of membrane-bound rod phosphodiesterase by the rod phosphodiesterase recombinant delta subunit. *J. Bio. Chem.* **271**, 24036-24047.

Forrester, J.V., Dick, A.D., McMenamin, P. and Lee, W.R. (1995). *The eye: basic sciences in practice.* WB Saunders Company Ltd, London, UK.

Fossarello, M., Serra, A., Mansfield, D., Wright, A., Loudianos, J., Piratsu, M. and Orzalesi, N. (1993). Genetic and epidemiological study of autosomal dominant retinitis pigmentosa (ADRP) and autosomal recessive (ARRP) retinitis pigmentosa in Sardinia. *In*: J.G. Hollyfield, R.E. Anderson and N. Orzalesi (eds.) *Retinal Degeneration*, pp.79-90, Plenum Press, New York, USA.

Francke, U., Ochs, H.D., De Martinville, B., Giacalone, J., Lindgren, V., Disteche, C., Pagon, R.A., Hofker, M.H., van Ommen, G.J.B., Pearson, P.L. and Wedgwood, R.J. (1985). Minor Xp21 chromosome deletion in a male associated with expression of Duchenne muscular dystrophy, chronic granulomatous disease, retinitis pigmentosa and McLeod syndrome. *Am. J. Hum. Genet.*, **37**, 250-267.

Freund, C.L., Gregory-Evans, C.Y., Furukawa, T., Papaioannou, M., Looser, J., Ploder, L., Bellingham, J., Ng, D., Herbrick, J.A., Duncan, A., Scherer, S.W., Tsui, L.C., Loutradis-Anagnostou, A., Jacobson, S.G., Cepko, C.L., Bhattacharya, S.S. and McInnes, R.R. (1997). Cone-rod dystrophy due to mutations in a novel photoreceptor-specific homeobox gene (Crx) essential for maintenance of the photoreceptor. *Cell*, **91**, 543-553.

Freund, C.L., Wang, Q.-L., Chen, S., Muskat, B.L., Wiles, C.D., Sheffield, V.C., Jacobson, S.G., McInnes, R.R., Zack, D.J. and Stone, E.M. (1998). De Novo mutations in the Crx homeobox gene associated with Leber congenital amaurosis. (Letter) *Nature Genet.*, **18**, 311-312.

Fuchs, S., Nakazawa, M., Maw, M., Tamai, M., Oguchi, Y. and Gal, A. (1995). A homozygous 1-base pair deletion in the arrestin gene is a frequent cause of Oguchi disease in Japanese. *Nature Genet.*, **10**, 360-362.

Fukuda, M., Aruga, J., Niinobe, M., Aimoto, S. and Mikoshiba, K. (1994). Inositol-1,3,4,5-tetrakisphosphate binding to C_2B domain of IP4BP/ synaptotagmin II. *J. Biol. Chem.*, **269**, 29206-29211.

Furukawa, T., Morrow, E.M. and Cepko, C.L. (1997). Crx, a novel Otx-like homeobox gene, shows photoreceptor-specific expression and regulates photoreceptor differentiation. *Cell*, **91**, 531-541.

Gal, A., Li, Y., Thompson, D.A., Weir, J., Orth, U., Jacobson, S.G., Apfelstedt-Sylla, E. and Vollrath, D. (2000). Mutations in MERTK, the human orthologue of the RCS rat retinal dystrophy gene, cause retinitis pigmentosa. *Nature Genet.*, **26**, 270-1.

Gal, A., Orth, U., Baehr, W., Schwinger, E. and Rosenberg, T. (1994). Heterozygous missense mutation in the rod cGMP phosphodiesterase beta-subunit gene in autosomal dominant stationary night blindness. *Nature Genet.*, **7**, 64-68.

Gardiner-Garden, M. and Frommer, M. (1987). CpG islands in vertebrate genomes. *J.Mol.Biol.* **196**, 261-276.

Geppert, M., Goda, Y., Hammer, R.E., Li, C., Rosahl, T.W., Stevens, C.F. and Südhof, T.C. (1994). Synaptotagmin I: a major Ca^{2+} sensor for transmitter release at a central synapse. *Cell*, **79**, 717-727.

Gieser, L., Fujita, R., Goring, H.H.H., Ott, J., Hoffman, D.R., Cideciyan, A.V., Birch, D.G., Jacobson, S.G. and Swaroop, A. (1998). A novel locus (RP24) for X-linked retinitis pigmentosa maps to Xq26-27. *Am. J. Hum. Genet.*, **63**, 1439-1447.

Gilbert, S.F. (1997). *Developmental Biology*, Sinauer Associates Inc, Massachusetts, USA.

Giniger, E. and Ptashne, M. (1988). Cooperative DNA binding of the yeast transcriptional activator GAL4. *Proc. Nat. Acad. Sci.*, **85**, 382-386.

Giniger, E., Varnum, S.M. and Ptashne, M. (1985). Specific DNA binding GAL4, a positive regulatory protein of yeast. *Cell*, **40**, 767-774.

Goldberg, A.F.X. and Molday, R.S. (1996). Subunit composition of the peripherin/Rds-Rom-1 disk rim complex from rod photoreceptors: hydrodynamic evidence for a tetrameric quaternary structure. *Biochemistry*, **35**, 6144-6149.

Goldberg, J. (1998). Structural basis for activation of ARF GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching. *Cell*, **95,** 237–248.

Goud, B., Zahraoui, A., Tavitian, A. and Saraste, J. (1990). Small GTP-binding protein associated with Golgi cisternae. *Nature* **345**, 553-556.

Graber, J., Cantor, C., Mohr, S. and Smith, T. (1999). *In silico* detection of control signals: mRNA 3'-end-processing sequences in diverse species. *Proc. Nat. Acad. Sci.*, **96**, 14055-14060.

Gray, I.C., Nobile, C., Muresu, R., Ford, S. and Spurr, N.K. (1995). A 2.4-megabase physical map spanning the CYP2C gene cluster on chromosome 10q24. *Genomics*, **28**, 328-332.

Grondahl, J. (1986). Tapeto-retinal degeneration in four Norwegian counties: II. Diagnostic evaluation of 407 relatives and genetic evaluation of 87 families. *Clin. Genet.*, **29**, 17-41.

Grondahl, J. (1987a). Autosomal recessive inheritance in "senile" retinitis pigmentosa. *Acta Ophthalmol.*, **65**, 231-236.

Grondahl, J. (1987b). Estimation of prognosis and prevalence of retinitis pigmentosa and Usher syndrome in Norway. *Clin. Genet.*, **31**, 255-264.

Gu, S., Lennon, A., Li, Y., Lorenz, B., Fossarello, M., North, M., Gal, A. and Wright, A. (1998). Tubby-like protein-1 mutations in autosomal recessive retinitis pigmentosa. *Lancet*, **351**, 1103-4.

Gu, S., Thompson, D.A., Srikumari, C.R.S., Lorenz, B., Finckh, U., Nicoletti, A., Murthy, K.R., Rathmann, M., Kumaramanickavel, G., Denton, M.J. and Gal, A. (1997). Mutations in RPE65 cause autosomal recessive childhood-onset severe retinal dystrophy. *Nature Genet.*, **17**, 194-197.

Guo, W., Roth, D., Walch-Solimena, C. and Novick, P. (1999). The exocyst is an effector for Sec4p, targeting secretory vesicles to sites of exocytosis. *EMBO J.* **18**, 1071-1080.

Guo, W., Sacher, M., Barrowman, J., Ferro-Novick, S. and Novick, P. (2000). Protein complexes in transport vesicle targeting. *Trends Cell Biol.* **10**, 251-5.

Gusella, J.F., Wexler, N.S., Conneally, P.M., Naylor, S.L., Anderson, M.A., Tanzi, R.E., Watkins, P.C., Ottina, K., Wallace, M.R., Sakaguchi, A.Y., Young, A.B., SHoulson, I., Bouilla, E. and Martin, J.B. (1983). A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* **306**, 234-8.

Guthrie, C. and Fink, G.R. (1991). Guide to yeast genetics and molecular biology. *In*: *Methods in enzymology* **194**, pp1-932, Academic Press, San Diego, USA.

Gyapay, G., Schmitt, K., Fizames, C., Jones, H., VegaCzarny, N., Spillett, D., Muselet, D., PrudHomme, J.F., Dib, C., Auffray, C., Morissette, J., Weissenbach, J. and Goodfellow, P.N. (1996). A radiation hybrid map of the human genome. *Hum. Mol. Genet.*, **5**, 339-346.

Hagstrom, S.A., Duyao, M., North, M.A. and Li, T. (1999). Retinal degeneration in tulp1-/- mice: vesicular accumulation in the interphotoreceptor matrix. *Invest. Ophthalmol. Vis. Sci.*, **40**, 2795-802.

Hagstrom, S.A., North, M.A., Nishina, P.M., Berson, E.L. and Dryja, T.P. (1998). Recessive mutations in the gene encoding the tubby-like protein TULP1 in patients with retinitis pigmentosa. *Nature Genet.*, **18**, 174-176.

Haider, N.B., Jacobson, S.G., Cideciyan, A.V., Swiderski, R., Streb, L.M., Searby, C., Beck, G., Hockey, R., Hanna, D.B., Gorman, S., Duhl, D., Carmi, R., Bennett, J., Weleber, R.G., Fishman, G.A., Wright, A.F., Stone, E.M. and Sheffield, V.C. (2000). Mutation of a nuclear receptor gene, NR2E3, causes enhanced S cone syndrome, a disorder of retinal cell fate. *Nature Genet.*, **24**, 127-131.

Haim, M. (1992). Prevalence of retinitis pigmentosa and allied disorders in Denmark: 3. Hereditary pattern. *Acta Ophthalmol.*, **70**, 615-624.

Hamel, C.P., Jenkins, N.A., Gilbert, D.J., Copeland, N.G. and Redmond, T.M. (1994). The gene for the retinal pigment epithelium-specific protein RPE65 is localized to human 1p31 and mouse 3. *Genomics*, **20**, 509-512.

Hamel, C.P., Tsilou, E., Pfeffer, B.A., Hooks, J.J., Detrick, B. and Redmond, T.M. (1993). Molecular cloning and expression of RPE65, a novel retinal pigment epithelium-specific microsomal protein that is post-transcriptionally regulated *in vitro*. *J. Biol. Chem.*, **268**, 15751-15757.

Hardcastle, A.J., Thiselton, D.L., van Maldergem, L., Saha, B.K., Jay, M., Plant, C., Taylor, R., Bird, A.C. and Bhattacharya, S. (1999). Mutations in the RP2 gene cause disease in 10% of families with familial X-linked retinitis pigmentosa assessed in this study. *Am. J. Hum. Genet.*, **64**, 1210-1215.

Hardcastle, A.J., Thiselton, D.L., Zito, I., Ebenezer, N., Mah, T.S., Gorin, M.B. and Bhattacharya, S.S. (2000). Evidence for a new locus for X-linked retinitis pigmentosa (RP23). *Invest. Ophthalmol. Vis. Sci.*, **41**, 2080-6.

Harlow, E. and Lane, D. (1988). *Antibodies: A Laboratory Manual*, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York, USA.

Harper, J., Adami, G., Wei, N., Keyomarsi, K. and Elledge, S. (1993). The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell*, **75**, 805-816.

Hartl, D.L. and Jones, E.W. (1998). *In*: K.E. Davis (ed.) *DNA Cloning*, IRL Press Ltd, Oxford, UK.

Heckenlively, J.R. (1988a). The diagnosis and classification of retinitis pigmentosa. *In:* J.R. Heckenlively (ed.) *Retinitis Pigmentosa*, pp.6-24, J.B. Lippincott Company Company, Philadelphia, USA.

Heckenlively, J.R. (1988b). Clinical findings in retinitis pigmentosa. *In*: J.R. Heckenlively, (ed.) *Retinitis Pigmentosa*, pp.68-89, J.B. Lippincott Company, Philadelphia, USA.

Heckenlively, J.R. (1988c). Retinitis Pigmentosa. *In*: J.R. Heckenlively (ed.) *Retinitis Pigmentosa*, pp.1-5, J.B. Lippincott Company, Philadelphia, USA.

Heckenlively, J.R. (1988d). Autosomal dominant retinitis pigmentosa. *In*: J.R. Heckenlively (ed.) *Retinitis Pigmentosa*, pp.125-149, J.B. Lippincott Company, Philadelphia, USA.

Heckenlively, J.R. and Bird, A.C. (1988). Choroideremia. *In*: J.R. Heckenlively (ed.) *Retinitis Pigmentosa*, pp.176-187, J.B. Lippincott Company, Philadelphia, USA.

Heckenlively, J.R. and Daiger, S.P. (2000). Mutations in a new photoreceptor-pineal gene on 17p cause Leber congenital amaurosis. *Nature Genet.*, 24: 79-83, 2000.

Heckenlively, J.R., Boughman, J.A. and Friedman, L.H. (1988). Pedigree analysis. *In*: J.R. Heckenlively (ed.) *Retinitis Pigmentosa*, pp.14-24, J.B. Lippincott Company, Philadelphia, USA.

Henikoff, S. and Henikoff, J. (1991). Automated assembly of protein blocks for database searching. *Nucleic Acids Res.*, **19,** 6565-6572.

Hiles, I.D., Otsu, M., Volinia, S., Fry, M.J., Gout, I., Dhand, R., Panayotou, G., Ruiz-Larrea, F., Thompson, A., Totty, N.F., Hsuan, J.J., Courtneidge, S.A., Parker, P.J. and Waterfield, M.D. (1992). Phosphatidylinositol 3-kinase: structure and expression of the 110 kd catalytic subunit. *Cell*, **70**, 419-429.

Hillier, L., Lennon, G., Becker, M., Bonaldo, M., Chiapelli, B., Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing,

T., Schellenberg, K., Soares, M., Tan, F., Trevaskis, E., Underwood, K., Wohldman, P., Waterston, R., Wilson R. and M. Marra. (1996). Generation and analysis of 280,000 human expressed sequence tags. *Genome Research* **6**, 807-828.

Hirokawa, T., Boon-Chieng, S. and Mitaku, S. (1998). SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics*, **14**, 378-379.

Hoffman, C. and Winston, F. (1987). A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. *Gene*, **57**, 267-272.

Hofmann, K. and Stoffel, W. (1993). TMbase – A database of membrane spanning protein segments. *Biol. Chem.*, **347**, 166.

Hofmann, K., Bucher, P., Falquet L. and Bairoch, A. (1999). The PROSITE database, its status in 1999. *Nucleic Acids Res.*, **27**, 215-219.

Hoger, T.H., Zatloukal, K., Waizenegger, I. and Krohne, G. (1990). Characterization of a second highly conserved B-type lamin present in cells previously thought to contain only a single B-type lamin. *Chromosoma*, **99**, 379-90.

Hong, D., Pawlyk, B.S., Shang, J., Sandberg, M.A., Berson, E.L. and Li, T. (2000b). A retinitis pigmentosa GTPase regulator (RPGR) deficient mouse model for X-linked retinitis pigmentosa (RP3). *Proc. Nat. Acad. Sci.*, **97**, 3649-3654.

Hong, D., Yue, G., Adamain, M. and Li, T. (2000a). A retinitis pigmentosa GTPase regulator (RPGR) - interacting protein is stably associated with the photoreceptor ciliary axoneme and anchors RPGR to the connecting cilium. *J. Biol. Chem.* epub [ahead of print].

Hong, J.-X., Lee, F.-J.S., Patton, W.A., Lin, C.-Y., Moss, J. and Vaughan, M. (1998). Phospholipid- and GTP-dependent activation of cholera toxin and phospholipase D by human ADP-ribosylation factor-like protein 1 (HARL1). *J. Biol. Chem.*, **273**, 15872-15876.

Hope, I.A. and Struhl, K. (1986). Functional dissection of a eukaryotic transcription protein, GNN4 of yeast. *Cell*, **46**, 885-894.

Howes, K., Bronson, J.D., Dang, L.Y., Li, N., Zhang, K., Ruiz, C., Helekar ,B., Lee, M., Subbaraya, I., Kolb, H., Chen, J. and Baehr, W. (1998). Gene array and expression of mouse retina guanylate cyclase activating proteins 1 and 2. *Invest. Ophthalmol. Vis. Sci.*, **39**, 867-75.

Hu, D. (1987). Prevalence and mode of inheritance of major genetic eye diseases in China. *J. Med. Genet.*, **24**, 584-588.

Huang, S.H., Pittler, S.J., Huang, X., Oliveira, L., Berson, E.L. and Dryja, T.P. (1995). Autosomal recessive retinitis pigmentosa caused by mutations in the alpha subunit of rod cGMP phosphodiesterase. *Nature Genet.*, **11**, 468-471.

Hubbard, R. (1958). The thermal stability of rhodopsin and opsin. *J. Gen. Physiol.*, **42**, 259-80.

Huber, L.A., de Hoop, M.J., Dupree, P., Zerial, M., Simons, K. and Dotti, C. (1993b). Protein transport to the dendritic plasma membrane of cultured neurons is regulated by rab8p. *J. Cell Biol.* **123**, 47-55.

Huber, L.A., Dupree, P. and Dotti, C.G. (1995). A deficiency of the small GTPase rab8 inhibits membrane traffic in developing neurons. *Mol. Cell. Biol.* **15**, 918-24.

Huber, L.A., Pimplikar, S., Parton, R.G., Virta, H., Zerial, M. and Simons, K. (1993a). Rab8, a small GTPase involved in vesicular traffic between the TGN and the basolateral membrane. *J. Cell Biol.* **123**, 35-45.

Humphries, P., Kenna, P. and Farrar, G.J. (1992). On the molecular genetics of retinitis pigmentosa. *Science*, **256**, 804-808.

Hunter, D.G., Fishman, G.A. and Kretzer, F.L. (1988). Abnormal axonemes in X-linked retinitis pigmentosa. *Arch. Ophthalmol.*, **106**, 362-368.

Hurwitz, R.L., Bunt-Milam, A.H., Chang, M.L. and Beavo, J.A. (1985). cGMP phosphodiesterase in rod and cone outer segments of the retina. *J. Bio. Chem.*, **260**, 568-573.

Hussels-Maumenee, I., Pierce, E.R., Bias, W.B. and Schleutermann, D.A. (1975). Linkage studies of typical retinitis pigmentosa and common markers. *Am. J. Hum. Genet.*, **27**, 505-508.

Hyatt, D., Snoddy, J., Schmoyer, D., Chen, G., Fischer, K., Parang, M., Vokler, I., Petrov, S., Locascio, P., Olman, V., Land, M., Shah, M. and Uberbacher, E. (2000). Improved analysis and annotation tools for whole-genome computational annotation and analysis: GRAIL-EXP genome analysis toolkit and related analysis tools. *Genome Sequencing & Biology Meeting*, May 2000.

Ikeda, S., Shiva, N., Ikeda, A., Smith, R.S., Nusinowitz, S., Yan, G., Lin, T.R., Chu, S., Heckenlively, J.R., North, M.A., Naggert, J.K., Nishina, P.M. and Duyao, M.P. (2000). Retinal degeneration but not obesity is observed in null mutants of the tubby-like protein 1 gene. *Hum. Molec. Genet.*, **9**, 155-163.

International Batten Disease Consortium (1995). Isolation of a novel gene underlying Batten disease, CLN3. *Cell*, **82**, 949-957.

Ish-Horowicz, D. and Burke, J.F. (1981). Rapid and efficient cosmid cloning. *Nucleic Acid Res.* **9**, 2989.

Jacobson, S.G., Cideciyan, A.V., Huang, Y., Hanna, D.B., Freund, C.L., Affatigato, L.M., Carr R.E., Zack D.J., Stone E.M. and McInnes R.R. (1998). Retinal degenerations with truncation mutations in the cone-rod homeobox (Crx) gene. *Invest. Ophthalmol. Vis. Sci.*, **39**, 2417-2426.

Jacobson, S.G., Cideciyan, A.V., Regunath, G., Rodriguez, F.J., Vandenburgh, K. Sheffield, V.C. and Stone, E.M. (1995). Nightblindness in Sorsby's fundus dystrophy is reversed by vitamin A. *Nature Genet.*, **11**, 27-32.

Jansen, G.A., Ofman, R., Ferdinandusse, S., Ijlst, L., Muijsers, A.O., Skjeldal, O.H., Stokke, O., Jakobs, C., Besley, G.T.N., Wraith, J.E. and Wanders, R.J.A. (1997). Refsum disease is caused by mutations in the phytanoyl-coa hydroxylase gene. *Nature Genet.*, **17**, 190-193.

Jasin, M., Regan, L. and Schimmel, P. (1984). Dispensable pieces of an aminoacyl tRNA synthetase which activate the catalytic site. *Cell* **36**, 1089-95.

Jasmin, B., Goud, B., Camus, G. and Cartaud, J. (1992). The low molecular weight guanosine triphosphate binding protein rab6p associates with distinct post-Golgi vesicles in *Torpedo marmorata* elecrolytes. *Neuroscience* **49,** 849-855.

Jay, M. (1982). On the heredity of retinitis pigmentosa. *Br. J. Ophthalmol.*, **66**, 405-416.

Jean, D., Ewan, K. and Gruss, P. (1998). Molecular regulators involved in vertebrate eye development. *Mech. Dev.*, **76**, 3-18.

Jiang, M., Pandey, S. and Fong, H.K.W. (1993). An opsin homologue in the retina and pigment epithelium. *Invest. Ophthal. Visual Sci.*, **34**, 3669-3678.

Jindeová, H. (1998). Vertebrate phototransduction: activation, recovery, and adaptation. *Physiol. Res.*, **47**, 155-168.

Johnston, M., Flick, J.S. and Pexton, T. (1994). Multiple mechanisms provide rapid and stringent glucose repression of *gal* gene expression in *Saccharomyces cerevisiae*. *Mol. Cell. Bio.*, **14**, 3834-3841.

Jones, S.M., Crosby, J.R., Salamero, J. and Howell, K.E. (1993). A cytosolic complex of p62 and rab6 associates with TGN38/41 and is involved in budding of exocytic vesicles from the tran-Golgi network. *J. Cell Biol.* **122**, 775-788.

Kahana, J.A. and Cleveland, D.W. (1999). Beyond nuclear transport: Ran-GTP as a determinant of spindle assembly. *J. Cell Biol.*, **146**, 1205-1209.

Kahn, R.A. (1996). The ARF subfamily. *In*: M. Zerial and L. Huber (eds.) *Guidebook to the small GTPases*, pp.429-433, Oxford Univserity Press/Sambrook and Tooze Publications, Oxford, UK.

Kaplan, J., Bonneau, D., Frezal, J., Munnich, A. and Dufier, J.L. (1990). Clinical and genetic heterogeneity in retinitis pigmentosa. *Hum. Genet.*, **85,** 635-642.

Kaplan, J., Pelet, A., Martin, C., Delrieu, O., Aymé, S., Bonneau, D., Briard, M.L., Hanauer, A., Larget-Piet, L., Michel-Awad, A., Plauchu, H., Dufier, J.L., Frezal, J. and Munnich, A. (1992). Phenotype-genotype correlations in X linked retinitis pigmentosa. *J. Med. Genet.*, **29**, 615-623.

Keegan, L., Gill, G. and Ptashne, M. (1986). Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. *Science*, **231**, 699-704.

Keen, T. J. and Inglehearn, C.F. (1996). Mutations and polymorphisms in the human peripherin-RDS gene and their involvement in inherited retinal degeneration. *Hum. Mutat.*, **8**, 297-303.

Kelsell, D.P., Rooke, L., Warne, D., Bouzyk, M., Cullin, L., Cox, S., West, L., Povey, S. and Spurr, N.K. (1995). Development of a panel of monochromosomal somatic cell hybrids for rapid gene mapping. *Ann. Hum. Genet.*, **59**, 233-241.

Kelsell, R.E., Gregory-Evans, K., Payne, A.M., Perrault, I., Kaplan, J., Yang, R.-B., Garbers, D.L., Bird, A.C., Moore, A.T. and Hunt, D.M. (1998). Mutations in the retinal guanylate cyclase (RETGC-1) gene in dominant cone-rod dystrophy. *Hum. Molec. Genet.*, **7**, 1179-1184.

Khani, S.C., Abitbol, M., Yamamoto, S., Maravic-Magovcevic, I. and Dryja, T.P. (1996). Characterization and chromosomal localization of the gene for human rhodopsin kinase. *Genomics*, **35**, 571-576.

Kirschner, R., Rosenberg, T., Schultz-Heienbrok, R., Lenzner, S., Feil, S., Roepman, R., Cremers, F.P.M., Ropers, H.-H. and Berger, W. (1999). RPGR transcription studies in mouse and human tissues reveal a retina-specific isoform that is disrupted in a patient with X-linked retinitis pigmentosa. *Hum. Molec. Genet.*, **8**, 1571-1578.

Klebe, C., Prinz, H., Wittinghofer, A. and Goody, R.S. (1995).The kinetic mechanism of Ran--nucleotide exchange catalyzed by RCC1.*Biochemistry* **34**, 12543-52.

Kleyn, T.J., Fan, W., Kovats, S.G., Lee, J.J., Pulido, J.C., Wu, Y,. Berkmeier, L.R., Misumi, D.J., Holmgren, L., Charlat, O., Woolf, E.A., Tayber, O., Brody, T., Shu, P., Hawkins, F., Kennedy, B., Baldini, L., Ebeling, C., Alperin, J.D., Deeds, J., Lakey, N.D., Culpepper, J., Chen, H., Glücksmann-Kuis, M.A., Carlson, G.A., Duyk, G.M. and Moore, K.J. (1996). Identification and characterization of the mouse obesity gene tubby: a member of a novel gene family. *Cell*, **85**, 281-290.

Knopf, J.L., Lee, M.H., Sultzman, L.A., Kriz, R.W., Loomis, C.R., Hewick, R.M. and Bell, R.M. (1986). Cloning and expression of multiple protein kinase C cDNAs. *Cell*, **46**, 491-502.

Kohl, S., Marx, T., Giddings, I., Jägle, H., Jacobson, S.G., Apfelstedt-Sylla, E., Zrenner, E., Sharpe, L.T. and Wissinger, B. (1998). Total colourblindness is caused by mutations in the gene encoding the alpha-subunit of the cone photoreceptor cGMP-gated cation channel. *Nature Genet.*, **19**, 257-259.

Köhler, M., Haller, H. and Hartmann, E. (1999). Nuclear protein transport pathways. *Exp. Nephrol.*, **7**, 290-4.

Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger-RNAs. *Nucleic Acids Res.*, **15**, 8125-8148.

Kramer, H. and Phistry, M. (1999). Genetic analysis of hook, a gene required for endocytic trafficking in drosophila. *Genetics*, **151**, 675-84.

Kumar, R., Chen, S., Scheurer, D., Wang, Q., Duh, E., Sung, C., Rehemtulla, A., Swaroop, A., Adler, R. and Zack, D. (1996). The bZIP transcription factor Nrl stimulates rhodopsin promoter activity in primary retinal cell cultures. *J. Biol. Chem.*, **271**, 29612-29618.

Kundu, T.K. and Rao, M.R. (1999). CpG islands in chromatin organization and gene expression. *J. Biochem.*, **125**, 217-222.

Lambright, D.G., Noel, J., Hamm, H.E. and Sigler, P.B. (1994). Structural determinants for the activation of the alpha-subunit of a heterotrimeric G-protein. *Nature*, **369**, 621-628.

Lambright, D.G., Sondek, J., Bohm, A., Skiba, N.P., Hamm, H.E. and Sigler, P.B. (1996). The 2.0 Å crystal structure of a heterotrimeric G protein. *Nature*, **379**, 311-319.

Langton, K.P., McKie, N., Curtis, A., Goodship, J.A., Bond, P.M., Barker, M.D. and Clarke, M. (2000). A novel tissue inhibitor of metalloproteinases-3 mutation reveals a common molecular phenotype in Sorsby's fundus dystrophy. *J. Biol. Chem.*, **275**, 27027-31.

Lee, R.M., Rossman, C.M., Forrest, J.B. and Newhouse, M.T. (1983). No evidence for pathological ciliary structure in retinitis pigmentosa. *J. Clin. Dysmorphol.* 1, 29-30.

Li, C., Ullrich, B., Zhang, J.Z., Anderson, R.G.W., Brose, N. and Südhof, T.C. (1995). Ca(2+)-dependent and -independent activities of neural and non-neural synaptotagmins. *Nature*, **375**, 594-599.

Li, J. and Herskowitz, I. (1993). Isolation of *ORC6*, a component of the yeast origin recognition complex by a one-hybrid system. *Science*, **262**, 1870-1874.

Li, T., Sandberg, M.A., Pawlyk, B.S., Rosner, B., Hayes, K.C., Dryja, T.P. and Berson, E.L. (1998). Effect of vitamin A supplementation on rhodopsin mutants threonine-17-to-methionine and proline-347-to-serine in transgenic mice and in cell cultures. *Proc. Nat. Acad. Sci.*, **95**, 11933-11938.

Li, T., Snyder, W.K., Olsson, J. E. and Dryja, T.P. (1996). Transgenic mice carrying the dominant rhodopsin mutation P347S: evidence for defective vectorial transport of rhodopsin to the outer segments. *Proc. Nat. Acad. Sci.*, **93**, 14176-14181.

Linari, M., Ueffing, M., Manson, F., Wright, A., Meitinger, T. and Becker, J. (1999). The retinitis pigmentosa GTPase regulator, RPGR, interacts with the delta subunit of rod cyclic GMP phosphodiesterase. *Proc. Nat. Acad. Sci.*, **96**, 1315-1320.

Liu, X., Garriga, P. and Khorana, H.G. (1996). Structure and function in rhodopsin: correct folding and misfolding in two point mutants in the intradiscal domain of rhodopsin identified in retinitis pigmentosa. *Proc. Nat. Acad. Sci.*, **93**, 4554-4559.

Liu, X., Vansant, G., Udovichenko, I.P., Wolfrum, U. and Williams, D.S. (1997). Myosin VIIa, the product of the Usher 1B syndrome gene, is concentrated in the connecting cilia of photoreceptor cells. *Cell Motil. Cytoskel.*, **37**, 240-252.

Lotery, A.J., Jacobson, S.G., Fishman, G.A., Weleber, R.G., Fulton, A.B., Namperumalsamy, P., Heon, E., Levin, A.V., Grover, S., Rosenow, J.R., Kopp, K.K., Sheffield, V.C. and Stone, E.M. (2001). Mutations in the CRB1 gene cause Leber congenital amaurosis. *Arch. Opthalmol.* **119**, 415-20.

Lounsbury, K.M., Richards, S.A., Carey, K.L. and Macara, I.G. (1996). Mutations within the Ran/TC4 GTPase. *J. Biol. Chem.*, **271**, 32834-32841.

Luban, J. and Goff, S.P. (1995). The yeast two-hybrid system for studying protein-protein interactions. *Curr. Opin. Biotechnol.*, **6**, 59-64.

Lupas, A. (1996). Coiled-coils: new structures and new functions. *Trends Biochem. Sci.*, **21**, 375-382.

Lupas, A., van Dyke, M. and Stock, J. (1991). Predicting coiled-coils from protein sequences. *Science*, **252**, 1162-1164.

Lyness, A.L., Ernst, W., Quinlan, M.P., Clover, G.M., Arden, G.B., Carter, R.M., Bird, A.C. and Parker, J.A. (1985). A clinical, psychological and electroretinographic survey of patients with autosomal dominant retinitis pigmentosa. *Br. J. Ophthalmol.*, **69**, 326-339.

Lyon, M.F. (1972). X-chromosome inactivation and developmental patterns in mammals. *Biol. Rev.*, **47**, 1-35.

Ma, J. and Ptashne, M. (1987). A new class of yeast transcriptional activators. *Cell*, **51**, 113-119.

Macrae, W. (1982). Retinitis pigmentosa in Ontario – a survey. *Birth Defects: Original Article Series*, **18**, 175-185.

Mahadevan, S. and Struhl, K. (1990). Tc, an unusual promoter element required for constitutive transcription of the yeast HIS3 gene. *Mol. Cell. Biol.*, **10**, 4447-55.

Marlhens, F., Bareil, C., Griffoin, J.-M., Zrenner, E., Amalric, P., Eliaou, C., Liu, S.-Y., Harris, E., Redmond, T.M., Arnaud, B., Claustres, M. and Hamel, C.P. (1997). Mutations in RPE65 cause Leber's congenital amaurosis. (Letter) *Nature Genet.*, **17**, 139-141.

Marquardt, A., Stöhr, H., Passmore, L.A., Krämer, F., Rivera, A., Weber and B.H.F. (1998). Mutations in a novel gene, VMD2, encoding a protein of unknown properties cause juvenile-onset vitelliform macular dystrophy (Best's disease). *Hum. Mol.Genet.* **7**, 1517-1525.

Martinez-Mir, A., Paloma, E., Allikmets, R., Ayoso, C., Del Rio, T., Dean, M., Vilageliu, L., Gonzalez-Duarte, R. and Balcells, S. (1998). Retinitis pigmentosa caused by a homozygous mutation in the Stargardt disease gene ABCR. *Nature Genet.*, **18**, 11-12.

Martinez, O., Schimdt, A., Salamero. J., Hoflack, B., Roa, M., and Goud, B. (1994). The small GTP-binding protein rab6 functions in intra-Golgi transport. *J. Cell Biol.* **127,** 1575–1588.

Marzesco, A., Galli, T., Louvard, D. and Zahroui, A. (1998). The rod cGMP phosphodiesterase delta subunit dissociates the small GTPase Rab13 from membranes. *J. Biol. Chem.*, **273**, 22340-22345.

Massof, R.W. and Finkelstein, D. (1981). Two forms of autosomal dominant primary retinitis pigmentosa. *Doc. Ophthalmol.*, **51**, 289-346.

Masuda, K., Xu, Z.J., Takahashi, S., Ito, A., Ono, M., Nomura, K. and Inoue, M. (1997). Peripheral framework of carrot cell nucleus contains a novel protein predicted to exhibit a long alpha-helical domain. *Exp. Cell. Res.*, **232**, 173-81.

Matynia, A., Mueller, U., Ong, N., Demeter, J., Granger, A.L., Hinata, K. and Sazer, S. (1998). Isolation and characterization of fission yeast sns mutants defective at the mitosis-to-interphase transition. *Genetics*, **148**, 1799-1811.

Maw, M.A., Corbeil, D., Koch, J., Hellwig, A., Wilson-Wheeler, J.C., Bridges, R.J., Kumaramanickavel, G., John, S., Nancarrow, D., Roper, K., Weigmann, A., Huttner, W. B. and Denton, M.J. (2000). A frameshift mutation in prominin (mouse)-like 1 causes human retinal degeneration. *Hum. Molec. Genet.*, **9**, 27-34.

Maw, M.A., Kennedy, B., Knight, A., Bridges, R., Roth, K.E., Mani, E.J., Mukkadan, J.K., Nancarrow, D., Crabb, J.W. and Denton, M.J. (1997). Mutation of the gene encoding cellular retinaldehyde-binding protein in autosomal recessive retinitis pigmentosa. *Nature Genet.*, **17**, 198-200.

McGee, T.L., Adams, S.M., Grimsby, J.L., Hong, D-H., Li, T., Andreasson, S., Berson, E.L., Dryja, T.P. (2001). Null mutations in the RPGR-interacting protein gene (*RPGRIP*) in patients with Leber congenital amaurosis. *Invest. Ophthalmol. Vis. Sci.***42**, B607.

McGuire, R.E., Sullivan, L.S., Blanton, S.H., Church, M.W., Heckenlively, J.R. and Daiger, S.P. (1995). X-linked dominant cone-rod degeneration: linkage mapping of a new locus for retinitis pigmentosa (RP15) to Xp22.13-p22.11. *Am. J. Hum. Genet.*, **57**, 87-94.

McLaughlin, M.E., Ehrhart, T.L., Berson, E.L. and Dryja, T.P. (1995). Mutation spectrum of the gene encoding the beta subunit of rod phosphodiesterase among patients with autosomal recessive retinitis pigmentosa. *Proc. Nat. Acad. Sci.*, **92**, 3249-3253.

McMahon, H.T. (1999). Endocytosis: an assembly protein for clathrin cages. *Curr. Biol.*, **9**, R332–R335.

Mears, A.J., Hiriyanna, S., Vervoort, R., Yashar, B., Gieser, L., Fahrner, S., Daiger, S.P., Heckenlively, J.R., Sieving, P.A., Wright, A.F. and Swaroop, A. (2000). Remapping of the RP15 locus for X-linked cone-rod degeneration to Xp11.4-p21.1, and identification of a De Novo insertion in the RPGR exon ORF15. *Am. J. Hum. Genet.*, **67**, 1000-1003.

Meindl, A., Carvalho, M.R., Herrmann, K., Lorenz, B., Achatz, H., Lorenz, B., Apfelstedt-Sylla, E., Wittwer, B., Ross, M. and Meitinger, T. (1995). A gene (SRPX) encoding a sushi-repeat-containing protein is deleted in patients with X-linked retinitis pigmentosa. *Hum. Mol. Genet.*, **4**, 2339-46.

Meindl, A., Dry, K., Herrmann, K., Manson, F., Ciccodicola, A., Edgar, A., Carvalho, M.R., Achatz, H., Hellebrand, H., Lennon, A, Migliaccio, C., Porter, K., Zrenner, E., Bird, A., Jay, M., Lorenz, B., Wittwer, B., D'Urso, M., Meitinger, T. and Wright, A. (1996). A gene (RPGR) with homology to the RCC1 guanine nucleotide exchange factor is mutated in X-linked retinitis pigmentosa (RP3). *Nature Genet.*, **13**, 35-42.

Menozzi, P., Piazza, A. and Cavalli-Sforza, L. (1978). Synthetic maps of human gene frequencies in Europeans. *Science*, **201**, 786-792.

Merin, S. and Auerbach, E. (1976). Retinitis pigmentosa. *Surv. Ophthalmol.*, **20**, 303-346.

Miano, M.G., Testa, F., Strazzullo, M., Trujillo, M., De Bernardo, C., Grammatico, B., Simonelli, F., Mangino, M., Torrente, I., Ruberto, G., Beneyto, M., Antinolo, G., Rinaldi, E., Danesino, C., Ventruto, V., D'Urso, M., Ayuso, C., Baiget, M. and Ciccodicola, A. (1999). Mutation analysis of the RPGR gene reveals novel mutations in south European patients with X-linked retinitis pigmentosa. *Eur. J. Hum. Genet.*, **7**, 687-94.

Mihalik, S.J., Morrell, J.C., Kim. D., Sacksteder, K.A., Watkins, P.A. and Gould, S.J. (1997). Identification of PAHX, a Refsum disease gene. *Nature Genet.*, **17**, 185-189.

Milam, A.H. and Li, Z. (1995). Retinal pathology in retinitis pigmentosa: Considerations for therapy. *In*: R.E. Anderson, M.M. LaVail and J.G. Hollyfield (eds.) *Degenerative Diseases of the Retina*, pp.275-284, Plenum Press, New York, USA.

Milano, J. and Strayer, D.S. (1998). Effects of overexpression of Ran/TC4 on mammalian cells *in vitro*. *Exp. Cell Res.*, **239**, 31-39.

Min, K.C., Zvyaga, T.A., Cypess, A.M. and Sakmar, T.P. (1993). Characterization of mutant

rhodopsins responsible for autosomal dominant retinitis pigmentosa: mutations on the cytoplasmic surface affect transducin activation. *J. Biol. Chem.*, **268**, 9400-9404.

Molday, L.L., Rabin, A.R. and Molday, R.S. (2000). ABCR expression in foveal cone photoreceptors and its role in Stargardt macular dystrophy. *Nature Genet.*, **25**, 257-8.

Molday, R.S. (1998). Photoreceptor membrane proteins, phototransduction, and retinal degenerative diseases – the Friedenwald lecture. *Invest. Ophthalmol. Vis. Sci.*, **39**, 2493-2513.

Moore, M.S. (1998). Ran and nuclear transport. *J. Biol. Chem.*, **273**, 22857-22860.

Morimura, H., Berson, E.L. and Dryja, T.P. (1999). Recessive mutations in the RLBP1 gene encoding cellular retinaldehyde-binding protein in a form of retinitis punctata albescens. *Invest. Ophthalmol. Vis. Sci.*, **40**, 1000-1004.

Morimura, H., Fishman, G.A., Grover, S.A., Fulton, A.B., Berson, E.L. and Dryja, T.P. (1998). Mutations in the RPE65 gene in patients with autosomal recessive retinitis pigmentosa or Leber congenital amaurosis. *Proc. Nat. Acad. Sci.*, **95**, 3088-3093.

Morimura, H., Saindelle-Ribeaudeau, F., Berson, E.L. and Dryja, T. P. (1999). Mutations in RGR, encoding a light-sensitive opsin homologue, in patients with retinitis pigmentosa. (Letter) *Nature Genet.*, **23**, 393-394.

Morinaga, N., Moss, J. and Vaughan, M. (1997). Cloning and expression of a cDNA encoding a bovine brain brefeldin A-sensitive guanine nucleotide-exchange protein for ADP-ribosylation factor. *Proc. Nat. Acad. Sci.*, **94**, 12926-12931.

Morrow, E.M., Furukawa, T. and Cepko, C.L. (1998). Vertebrate photoreceptor cell development and disease. *Trends Cell. Biol.*, **8**, 353-358.

Mount, S. (1982). A catalogue of splice junction sequences. *Nucleic Acids Res.*, **10**, 459-472.

Murvai, J., Vlahovicek, K., Barta, E., Cataletto, B. and Pongor, S. (2000). The SBASE protein domain library, release 7.0: a collection of annotated protein sequence segments. *Nucleic Acids Res.*, **28**, 260-262.

Musarella, M.A., Anson-Cartwright, L., Leal, S.M., Gilbert, L.D., Worton, R.G., Fishman, G.A. and Ott, J. (1990). Multipoint linkage analysis and heterogeneity testing in 20 X-linked retinitis pigmentosa families. *Genomics*, **8**, 286-296.

Musarella, M.A., Anson-Cartwright, L., McDowell, C., Burghes, A., Coulson, S.E., Worton, R.G. and Rommens, J.M. (1991). Physical mapping at a potential X-linked retinitis pigmentosa locus (RP3) by pulsed field gel electrophoresis. *Genomics*, **11**, 263-272.

Nagase, T., Ishikawa, K., Suyama, M., Kikuno, R., Hirosawa, M., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N. and Ohara, O. (1999). Prediction of the coding sequences of unidentified human genes. XIII. The complete sequences of 100 new cDNA clones from brain which code for large proteins *in vitro*. *DNA Res.*, **1**, 63-70.

Nakai, K. and Horton, P. (1999). PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem. Sci.*, **24**, 34-35.

Nakajima, H., Hirata, A., Ogawa, Y., Yonehara, T., Yoda, K. and Yamasaki, M. (1991). A cytoskeleton-related gene, USO1, is required for intracellular protein transport in *Saccharomyces cerevisiae*. *J. Cell Biol.*, **113**, 245-60.

Nakazawa, M., Wada, Y. and Tamai, M. (1998). Arrestin gene mutations in autosomal recessive retinitis pigmentosa. *Arch. Ophthalmol.*, **116**, 498-501.

Nalefski, E.A., McDonagh, T., Somers, W., Seehra, J., Falke, J.J. and Clark, J.D. (1998). Independent folding and ligand specificity of the C_2 calcium-dependent lipid binding domain of cytosolic phospholipase A2. *J. Biol. Chem.*, **273**, 1365-1372.

Nalefski, E.A., Slazas, M.M. and Falke, J.J. (1997). Ca^{2+}-signaling cycle of a membrane-docking C_2 domain. *Biochemistry*, **36**, 12011-12018.

Nalefski, E.A., Sultzman, L.A., Martin, D.M., Kriz, R.W., Towler, P.S., Knopf, J.L. and Clark, J.D. (1994). Delineation of two functionally distinct domains of cytosolic phospholipase A2, a regulatory Ca(2+)-dependent lipid-binding domain and a Ca(2+)-independent catalytic domain. *J. Biol. Chem.*, **269**, 18239-18249.

Nathans, J. and Hogness, D.S. (1984). Isolation and nucleotide sequence of the gene encoding human rhodopsin. *Proc. Nat. Acad. Sci.*, **81**, 4851-4855.

Nathans, J., Davenport, C.M., Maumenee, I.H., Lewis, R.A., Hejtmancik, J.F., Litt, M., Lovrien, E., Weleber, R., Bachynski, B., Zwas, F., Kligamen, R. and Fishman, G. (1989). Molecular genetics of human blue-cone monochromacy. *Science*, **245**, 831-838.

Nathans, J., Maumenee, I.H., Zrenner, E., Sadowski, B., Sharpe, L.T., Lewis, R.A., Hansen, E., Rosenberg, T., Schwartz, M., Heckenlively, J.R., Traboulsi, E., Klingaman, R., Bech-Hansen, N.T, Laroche, G.R., Pagon, R.A., Murphey, W.H. and Weleber, R.G. (1993). Genetic heterogeneity among blue cone monochromats. *Am. J. Hum. Genet.*, **53**, 987-1000.

Nathans, J., Piantanida, T.P., Eddy, R.L., Shows, T.B. and Hogness, D.S. (1986b). Molecular genetics of inherited variation in human colour vision. *Science*, **232**, 203-210.

Nathans, J., Thomas, D. and Hogness, D.S. (1986a). Molecular genetics of human color vision: the genes encoding blue, green, and red pigments. *Science*, **232**, 193-202.

Neer, E.J., Schmidt, C.J., Nambudripad, R. and Smith, T.F. (1994). The ancient regulatory-protein family of WD-repeat proteins. *Nature*, **371**, 297-300.

Nettleship, E. (1908). On retinitis pigmentosa and allied diseases. *Roy. Lond. Ophhalmol. Hosp. Rep.*, **17**, 333-426.

Nicolas, F.J., Zhang, C.M., Hughes, M., Goldberg, M.W., Watton, S.J. and Clarke, P.R. (1997). Xenopus Ran-binding protein 1: molecular interactions and effects on nuclear assembly in Xenopus egg extracts. *J. Cell Sci.*, **110**, 3019-3030.

Nishimoto, T. (1999). A new role of Ran GTPase. *Biochem. Biophys. Res. Comm.* **262**, 571-574.

Nishizuka, Y. (1988). The molecular heterogeneity of protein kinase C and its implications for cellular regulation. *Nature*, **334**, 661-665.

North, M.A., Naggert, J.K., Yan, Y., Noben-Trauth, K. and Nishina, P.M. (1997). Molecular characterization of TUB, TULP1, and TULP2, members of the novel tubby gene family and their possible relation to ocular diseases. *Proc. Nat. Acad. Sci.*, **94**, 3128-3133.

Novick, P., and Brennwald, P. (1993). Friends and family: the role of the Rab GTPases in vesicular traffic. *Cell* **75,** 597–601.

Novick, P., and Zerial, M. (1997). The diversity of Rab proteins in vesicle transport.

Novina, C.D. and Roy, A.L. (1996). Core promoters and transcriptional control. *Trends Genet.*, **12**, 351-354.

Nuoffer, C., and Balch, W. E. (1994). GTPases: multifunctional molecular switches regulating vesicular traffic. *Annu. Rev. Biochem.* **63**, 949–990.

Nussbaum, R.L., Lewis, R.A., Lesko, J.G. and Ferrell, R. (1985). Mapping X-linked ophthalmic diseases: II. Linkage relationship of X-linked retinitis pigmentosa to X chromosomal short arm markers. *Hum. Genet.*, **70,** 45-50.

Ohga, H., Suzuki, T., Fujiwara, H., Furutani, A. and Koga, H. (1991). A case of immotile ciliea syndrom accompanied by retinitis pigmentosa. *Nippon Ganka Gakkai Zasshi* **95**, 795-801.

Ohtsubo, M., Kai, R., Furuno, N., Sekiguchi, T., Sekiguchi, M., Hayashida, H., Kuma, K., Miyata, T., Fukushige, S., Murotsu, T., Matsubara, K. and Nishimoto T. (1987). Isolation and characterisation of the active cDNA of the human cell cycle gene (RCC1) involved in the regulation of onset of chromosome condensation. *Genes Dev.*, **1**, 585-593.

Olsson, J.E., Gordon, J.W., Pawlyk, B.S., Roof, D., Hayes, A., Molday, R.S., Mukai, S., Cowley, G.S., Berson, E.L. and Dryja, T.P. (1992). Transgenic mice with a rhodopsin mutation (Pro23His): a mouse model of retinitis pigmentosa. *Neuron*, **9**, 815-830.

Osterberg, G. (1935). Topography of the layer of rods and cones in the human retina. *Acta Ophthalmol.*, **6**, 1-103.

Ott, J., Bhattacharya, S., Chen, J.D., Denton, M.J., Donald, J., Dubay, C., Farrar, G.J., Fishman, G.A., Frey, D., Gal, A., Humphries, P., Jay, B., Jay, M., Litt, M., Machler, M., Musarella, M., Neugebauer, M., Nussbaum, R.L., Terwilliger, J.D., Weleber, R.G., Wirth, B., Wong, F., Worton, R.G. and Wright, A.F. (1990). Localizing multiple X-chromosome linked retinitis pigmentosa loci using multilocus homogeneity tests. *Proc. Nat.Acad. Sci.*, **87**, 701-704.

Ouahchi, K., Arita, M., Kayden, H., Hentati, F., Ben Hamida, M., Sokol, R., Arai, H., Inoue,K., Mandel, J.-L. and Koenig, M. (1995). Ataxia with isolated vitamin E deficiency is caused by mutations in the alpha-tocopherol transfer protein. *Nature Genet.*, **9**, 141-145.

Palczewski, K., Kumasaka, T., Hori, T., Behnke, C.A., Motoshima, H., Fox, B.A., Le Trong, I., Teller, D.C., Okada, T., Stenkamp, R.E., Yamamoto, M. and Miyano, M. (2000). Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**, 739-45.

Pauling, L. and Corey, R.B. (1953). Compound helical configurations of polypeptide chains: Structure of proteins of the - keratin type. *Nature*, **171**, 59-61.

Payne, A.M., Downes, S.M., Bessant, D.A.R., Taylor, R., Holder, G.E., Warren, M.J., Bird, A.C. and Bhattacharya, S.S. (1998). A mutation in guanylate cyclase activator 1A (GUCA1A) in an autosomal dominant cone dystrophy pedigree mapping to a new locus on chromosome 6p21.1. *Hum. Molec. Genet.*, **7**, 273-277.

Pearlman, J.T. (1979). Mathematical models of retinitis pigmentosa: a study of the rate of progress in the different genetic forms. *Trans. Am. Ophthalmol. Soc.*, **77**, 643-656.

Pepe, I.M. (1999). Rhodopsin and phototransduction. *J. Photochem. Photobiol. B: Biol.*, **48**, 1-10.

Peränen, J., Auvinen, P., Virta, H., Wepf, R. and Simons, K. (1996). Rab8 promotes polarized membrane transport through reorganization of actin and microtubules in fibroblasts. *J. Cell Biol.* **135**, 153-167.

Perin, M.S., Brose, N., Jahn, R. and Südhof, T.C. (1991). Domain structure of synaptotagmin (p65). *J. Biol. Chem.*, **266**, 623-629.

Perin, M.S., Fried, V.A., Mignery, G.A., Jahn, R. and Südhof, T.C. (1990). Phospholipid binding by a synaptic vesicle protein homologous to the regulatory region of protein kinase C. *Nature*, **345**, 260-263.

Perisic, O., Fong, S., Lynch, D.E., Bycroft, M. and Williams, R.L. (1998). Crystal structure of a calcium-phospholipid binding domain from cytosolic phospholipase A2. *J. Biol. Chem.*, **273**, 1596-1604.

Perrault, I., Rozet, J.M., Calvas, P., Gerber, S., Camuzat, A., Dollfus, H., Chatelin, S., Souied, E., Ghazi, I., Leowski, C., Bonnemaison, M., Le Paslier, D., Frezal, J., Dufier, J.-L., Pittler, S.J., Baehr, W., Wasmuth, J.J., McConnell, D.G., Champagne, M.S., Vantuinen, P., Ledbetter, D. and Davis, R.L. (1990). Molecular characterization of human and bovine rod photoreceptor cGMP phosphodiesterase alpha-subunit and chromosomal localization of the human gene. *Genomics*, **6**, 272-283.

Perrault, I., Rozet, J.M., Calvas, P., Gerber, S., Camuzat, A., Dollfus, H., Châtelin, S., Souied, E.V., Ghazi, I., Leowski, C., Bonnemaison, M., Le Paslier, D., Frézal, J., Dufier, J-L., Pittler, S., Munnich, A. and Kaplan, J. (1996). Retinal-specific guanylate cyclase gene mutations in Leber's congenital amaurosis. *Nat. Genet.* **14**, 461-464.

Perrault, I., Rozet, J.M., Gerber, S., Ghazi, I., Leowski, C., Ducroq, D., Souied, E., Dufier, J.L., Munnich, A. and Kaplan, J. (1999). Leber congenital amaurosis. *Mol. Genet. Metab.*, **68**, 200-8.

Persson, B. and Argos, P. (1994). Prediction of transmembrane segments in proteins utilising multiple sequence alignments. *J. Mol. Biol.*, **237**, 182-192.

Pfeffer, S. (1994). Rab GTPases: master regulators of membrane trafficking. *Curr. Opin. Cell Biol.* **6**, 522–526.

Piazza, A., van Loghem, E., de Lange, G., Curtoni, E.S., Ulizzi, L. and Terrenato, L. (1976). Immunoglobulin allotypes in Sardinia. *Am. J. Hum. Genet.*, **28**, 77-86.

Pierce, E.A., Quinn, T., Meehan, T., McGee, T.L., Berson, E.L. and Dryja, T.P. (1999). Mutations in a gene encoding a new oxygen-regulated photoreceptor protein cause dominant retinitis pigmentosa. *Nat. Genet.* **22**, 248-254.

Pittler, S.J., Baehr, W., Wasmuth, J.J., McConnell, D.G., Champagne, M.S., van Tuinen, P., Ledbetter, D. and Davis, R.L. (1990). Molecular characterization of human and bovine rod photoreceptor cGMP phosphodiesterase alpha-subunit and chromosomal localization of the human gene. Genomics **6**, 272-83.

Pittler, S., Munnich, A. and Kaplan, J. (1996). Retinal-specific guanylate cyclase gene mutations in Leber's congenital amaurosis. *Nature Genet.*, **14**, 461-464.

Plant, P.J., Yeger, H., Staub, O., Howard, P. and Rotin, D. (1997). The C_2 domain of the ubiquitin protein ligase Nedd4 mediates Ca^{2+}-dependent plasma membrane localization. *J. Biol. Chem.*, **272**, 32329-32336.

Popoli, M., Venegoni, A., Buffa, L., Racagni, G. (1997). Ca^{2+}/phospholipid-binding and syntaxin-binding of native synaptotagmin I. *Life Sci.* **61**, 711-21.

Qin, N., Pittler, S.J. and Baehr, W. (1992). *In vitro* isoprenylation and membrane association of mouse rod photoreceptor cGMP phosphodiesterase alpha and beta subunits expressed in bacteria. *J. Biol. Chem.*, **267**, 8458-8463.

Rattner, A., Smallwood, P.M. and Nathans, J. (2000). Identification and characterization of all-trans-retinol dehydrogenase from photoreceptor outer segments, the visual cycle enzyme that reduces all-trans-retinal to all-trans-retinol. *J. Biol. Chem.*, **275**, 11034-43.

Rattner, A., Sun, H. and Nathans, J. (1999). Molecular genetics of human retinal disease. *Ann. Rev. Genet.*, **33**, 89-131.

Redmond, T.M., Yu, S., Lee, E., Bok, D., Hamasaki, D., Chen, N., Goletz, P., Ma, J.-X., Crouch, R.K. and Pfeifer, K. (1998). RPE65 is necessary for production of 11-*cis* vitamin A in the retinal visual cycle. *Nature Genet.*, **20**, 344-351.

Reichel, E., Bruce, A.M., Sandberg, M.A. and Berson, E.L. (1989). An electroretinographic and molecular study of X-linked cone degeneration. *Am. J. Ophthalmol.*, **108**, 540-547.

Remé, C.E., Grimm, C., Hafezi, F., Marti, A. and Wenzel, A. (1998). Apoptotic cell death in retinal degenerations. *Prog. Ret. Eye Res.*, **17**, 443-464.

Renault, L., Nassar, N., Vetter, I., Becker, J., Klebe, C., Roth, M. and Wittinghoffer, A. (1998). The 1.7 Å crystal structure of the regulator of chromosome condensation (RCC1) reveals a seven-bladed propeller. *Nature*, **392**, 97-100.

Reuber, B.E., Germain-Lee, E., Collins, C.S., Morrell, J.C., Ameritunga, R., Moser, H.W., Valle, D. and Gould, S.J. (1997). Mutations in PEX1 are the most common cause of peroxisome biogenesis disorders. *Nature Genet.*, **17**, 445-448.

Rhee, S.G., Suh, P.-G., Ryu, S.-H. and Lee, S.Y. (1989). Studies of inositol phospholipid-specific phospholipase C. *Science*, **244**, 546-550

Rispoli, G. (1998). Calcium regulation of phototransduction in vertebrate rod outer segments. *J. Photochem. Photobiol. B: Biol.*, **44**, 1-20.

Rizo, J. and Südhof, T.C. (1998). C_2-domains, structure and function of a universal Ca^{2+}-binding domain. *J. Biol. Chem.*, **273**, 15879–15882.

Robinson, M.S., Watts, C. and Zerial, M. (1996). Membrane dynamics in endocytosis. *Cell*, **84**, 13-21.

Robinson, W.G., Kuwabara, T. and Bieri, J.G. (1982). The role of vitamin E and unsaturated fatty acids in the visual process. *Retina*, **2**, 263-281.

Roepman, R., van Duijnhoven, G., Rosenberg, T., Pinckers, A.J.L.G., Bleeker-Wagemakers, L.M., Bergen, A.A.B., Post, J., Beck, A., Reinhardt, R., Ropers, H.-H., Cremers, F.P.M. and Berger, W. (1996). Positional cloning of the gene for X-linked retinitis pigmentosa 3: homology with the guanine-nucleotide-exchange factor RCC1. *Hum. Molec. Genet.*, **5**, 1035-1041.

Roepman, R., Bernoud-Hubac, N., Schick, D.E., Maugeri, A., Berger, W., Ropers, H-H., Cremers, F.P.M. and Ferreira, P.A. (2000). The retinitis pigmentosa GTPase regulator (RPGR) interacts with novel transport-like proteins in the outer segments of rod photoreceptors. *Hum. Mol. Genet.* **9**, 2095-2105.

Rosa, J.L. and Barbacid, M. (1997). A giant protein that stimulates guanine nucleotide exchange on ARF1 and Rab proteins forms a cytosolic ternary complex with clathrin and Hsp70. *Oncogene*, **15**, 1-6.

Rosa, J.L., Casaroli-Marano, R.P., Buckler, A.J., Vilaró, S. and Barbacid, M. (1996a). P619, a giant protein related to the chromosome condensation regulator RCC1, stimulates guanine nucleotide exchange on ARF1 and rab proteins. *EMBO J.*, **15**, 4262-4273.

Rosa, J.L., Casaroli-Marano, R.P., Buckler, A.J., Vilaró, S. and Barbacid, M. (1996b). Erratum. *EMBO J.*, **15**, p51738.

Rost, B. and Sander, C. (1993). Prediction of protein structure at better than 70% accuracy. *J. Mol. Biol.*, **232**, 584-599.

Roth, M.G. (1999). Snapshots of ARF1: Implications for mechanisms of activation and inactivation. *Cell*, **97**, 149-152.

Rothman, J. E. (1994) Mechanisms of intracellular protein transport. *Nature* **372**, 55–63

Rothman, J.E. and Wieland, F.T. (1996). Protein sorting by transport vesicles. *Science*, **272**, 227–234.

Roux, A.F., Yuan, C.C., Rommens, J.M. and Musarella, M.A. (1993). Dinucleotide repeat polymorphism near the RP3 locus in Xp21 (DXS1110). *Hum. Mol. Genet.*, **2**, 821.

Ruiz-Avila, L., McLaughlin, S.K., Wildman, D., McKinnon, P.J., Robichon, A., Spickofsky, N. and Margolskee, R.F. (1995). Coupling of bitter receptor to phosphodiesterase through transducin in taste receptor cells. *Nature*, **376**, 80-85.

Ruohonen, L., Aalto, M. and Keranen, S. (1995). Modifications to the *ADH1* promoter of *Saccharomyces cerevisiae* for efficient production of heterologous proteins. *J. Biotech.*, **39**, 193-203.

Saari, J.C., Bredberg, D.L. and Noy, N. (1994). Control of substrate flow at a branch in the visual cycle. *Biochemistry*, **33**, 3106-3112.

Saari, J.C., Nawrot, M., Kennedy, B.N., Garwin, G.G., Hurley, J.B., Huang, J., Possin, D.E. and Crabb, J.W. (2001). Visual cycle impairment in cellular retinaldehyde binding protein (CRALBP) knockout mice results in delayed dark adaptation. *Neuron* **29**, 739-48.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.

Sandberg, M. A., Weigel DiFranco, C., Dryja, T. P. & Berson, E. L. (1995). Clinical expression correlates with location of rhodopsin mutation in dominant retinitis pigmentosa. *Invest. Ophthalmol. Visual Sci.* **36**, 1934-1942.

Satoh, A., Tokunaga, F., Kawamura, S., Ozaki, K. (1997). In situ inhibition of vesicle transport and protein processing in the dominant negative Rab1 mutant of Drosophila. *J. Cell Sci.* **110**, 2943-53.

Sauer, C.G., Gehrig, A., Warneke-Wittstock, R., Marquardt, A., Ewing, C.C., Gibson, A., Lorenz, B., Jurklies, B. and Weber, B.H.F. (1997). Positional cloning of the gene associated with X-linked juvenile retinoschisis. *Nature Genet.*, **17**, 164-170.

Sazer, S. and Dasso, M. (2000). The Ran decathlon: multiple roles of Ran. *J. Cell Sci.*, **113**, 1111-1118.

Schiavo, G., Gmachl, N.J., Stenbeck, G., Sollner, T.H. and Rothman, J.E. (1995). A possible docking and fusion particle for synaptic transmission. *Nature*, **378**, 733-736.

Schultz, J., Copley, R.R., Doerks, T., Ponting, C.P. and Bork, P. (2000). SMART: a web-based tool for the study of genetically mobile domains. *Nucleic Acids Res.* **28**, 231-4.

Schultz, J., Milpetz, F., Bork, P. and Ponting, C.P. (1998). SMART, a simple modular architecture research tool: identification of signaling domains. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5857-64.

Schürmann, A., Schmidt, M., Asmus, M., Bayer, S., Fliegert, F., Koling, S., Maßmann, S., Schilf, C., Subauste, M.C., Voß, M., Jakobs, K.H. and Joost H. (1999). The ADP-ribosylation factor (ARF)-related GTPase ARF-related protein binds to the ARF-specific guanine nucleotide exchange factor cytohesin and inhibits the ARF-dependent activation of phospholipase D. *J. Biol. Chem.*, **274**, 9744-9751.

Schwahn, U., Lenzner, S., Dong, J., Feil, S., Hinzmann, B., van Duijnhoven, G., Kirschner, R., Hemberger, M., Bergen, A.A.B., Rosenberg, T. and Pinckers, A. (1998). Positional cloning of the gene for X-linked retinitis pigmentosa 2. *Nature Genet.*, **19**, 327-332.

Scott, K. and Zuker, C. (1997). Lights out: deactivation of the phototransduction cascade. *Trends Biochem. Sci.*, **22**, 350-354.

Seabra, J.M., Ho, Y.K. and Anant, J.S. (1995). Deficient geranylgeranylation of Ram/Rab27 in choroideremia. *J. Biol. Chem.*, **270**, 24420-24427.

Seabra, M.C., Brown, M.S., Slaughter, C.A., Sudhof, T.C. and Goldstein, J.L. (1992). Purification of component A of Rab geranylgeranyl transferase: possible identity with the choroideremia gene product. *Cell*, **70**, 1049-1057.

Seeliger, M.W., Biesalski, H.K., Wissinger, B., Gollnick, H., Gielen, S., Frank, J., Beck, S. and Zrenner, E. (1999). Phenotype in retinol deficiency due to a hereditary defect in retinol binding protein synthesis. *Invest. Ophthalmol. Vis. Sci.*, **40**, 3-11.

Senapathy, P., Shapiro, M. and Harris, N. (1990). Splice junctions, branch point sites, and exons: sequence statistics, identification, and applications to genome projects. *Methods in Enzymol.*, **183**, 253-278.

Shao, X., Davletov, B.A., Sutton, R.B., Südhof, T.C. and Rizo, J. (1996). Bipartite Ca^{2+}-binding motif in C_2 domains of synaptotagmin and protein kinase C. *Science*, **273**, 248-251.

Sharp, D., Blinderman, L., Combs, K.A., Kienzle, B., Ricci, B., Wager-Smith, K., Gil, C.M., Turck, C.W. Bouma, M-E., Rader, D.J., Aggerbeck, L.P., Gregg, R.E., Gordon, D.A. and Wetterau, J.R. (1993). Cloning and gene defects in microsomal triglyceride transfer protein associated with abetalipoproteinaemia. *Nature*, **365**, 65-69.

Sheets, M.D., Ogg, S.C. and Wickens, M.P. (1990). Point mutations in AAUAAA and the poly (A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation *in vitro*. *Nucleic Acids Res.*, **18**, 5799-5805.

Sheng, Z.H., Yokoyama, C.T. and Catterall, W.A. (1997). Interaction of the synprint site of N-type Ca^{2+} channels with the C_2B domain of synaptotagmin I. *Proc. Nat. Acad. Sci.*, **94**, 5405-5410.

Shetty, K.M., Kurada, P. and O'Tousa, J.E. (1998). Rab6 Regulation of Rhodopsin Transport in *Drosophila*. *J. Biol. Chem.* 273, 20425-20430.

Shirataki, H., Kaibuchi, K., Sakoda, T., Kishida, S., Yamaguchi, T., Wada, K., Miyazaki, M. and Takai, Y. (1993). Rabphilin-3A, a putative target protein for smg p25A/rab3A p25 small GTP-binding protein related to synaptotagmin. *Mol. Cell. Biol.*, **13**, 2061-2068.

Shoulders, C.C., Brett, D.J., Bayliss, J.D., Narcisi, T.M.E., Jarmuz, A., Grantham, T.T., Leoni, P.R.D., Bhattacharya, S., Pease, R.J., Cullen, P.M., Levi, S., Byfield, P.G.H., Purkiss, P. and Scott, J. (1993). Abetalipoproteinemia is caused by defects of the gene encoding the 97 kDa subunit of a microsomal triglyceride transfer protein. *Hum. Mol. Genet.*, **2**, 2109-2116.

Siebert, P.D., Chenchik, A., Kellogg, D.E., Lukyanov, K.A. and Lukyanov, S.A. (1995). An improved method for walking in uncloned genomic DNA. *Nucleic Acids. Res.*, **23**, 1087-1088.

Silver, P.A., Keegan, L.P. and Ptashne, M. (1984). Amino terminus of the yeast GAL4 gene product is sufficient for nuclear localisation. *Proc. Nat. Acad. Sci.*, **81**, 5951-5955.

Simon, A., Hellman, U., Wernstedt, C. and Eriksson, U. (1995). The retinal pigment epithelial-specific 11-*cis* retinol dehydrogenase belongs to the family of short chain alcohol dehydrogenases. *J. Biol. Chem.*, **270**, 1107-1112.

Slavotinek, A.M., Stone, E.M., Mykytyn, K., Heckenlively, J.R., Green, J.S., Heon, E., Musarella, M.A., Parfrey, P.S., Sheffield, V.C. and Biesecker, L.G. (2000). Mutations in MKKS cause Bardet-Biedl syndrome. *Nature Genet.*, **26**, 15-16.

Sohocki, M.M., Bowne, S.J., Sullivan, L.S., Blackshaw, S., Cepko, C.L., Payne, A.M., Bhattacharya, S.S., Khaliq, S., Mehdi, S.Q., Birch, D.G., Harrison, W.R., Elder, F.F.B., Sohocki, M.M., Sullivan, L.S., Mintz-Hittner, H.A., Birch, D., Heckenlively, J.R., Freund, C.L., McInnes, R.R. and Daiger, S.P. (1998). A range of clinical phenotypes associated with mutations in Crx, a photoreceptor transcription-factor gene. *Am. J. Hum. Genet.*, **63**, 1307-1315.

Sohocki, M.M., Bowne, S.J., Sullivan, L.S., Blackshaw, S., Cepko, C.L., Payne, A.M., Bhattacharya, S.S., Khaliq, S., Qasim Mehdi, S., Birch, D.G., Harrison, W.R., Elder, F.F.B., Heckenlively, J.R. and Daiger, S.P. (2000). Mutations in a new photoreceptor-pineal gene on 17p cause Leber congenital amaurosis. *Nat. Genet.* **24,** 79-83.

Sollner, T., Bennett, M. K., Whiteheart, S. W., Scheller, R. H. and Rothman, J. E. (1993) A protein assembly-disassembly pathway *in vitro* that may correspond to sequential steps of synaptic vesicle docking, activation, and fusion. *Cell* **75**, 409–418.

Sollner, T.H. and Rothman, J.E. (1996). Molecular machinery mediating vesicle budding, docking and fusion. *Cell Struct. Funct.***21**, 407-12.

Somsel Rodman, J. and Wandinger-Ness, A.. (2000). Rab GTPases coordinate endocytosis. *J. Cell Sci.* **113**, 183-192.

Souied, E., Soubrane, G., Benlian, P., Coscas, G.J., Gerber, S., Munnich, A. and Kaplan, J. (1996). Retinitis punctata albescens associated with the Arg135Trp mutation in the rhodopsin gene. *Am. J. Ophthalmol.*, **121**, 19-25.

Sorsby, A. (1966). The incidence and causes of blindness in England and Wales 1948-1962. *Reports on Public Health and Medical Subjects* **No. 114**, HMSO, London.
Spencer, M., Detwiler, P.B. and Bunt-Milam, A.H. (1988). Distribution of membrane proteins in mechanically dissociated retinal rods. *Invest. Ophthalmol. Vis. Sci.* **29**, 1012-1020.

Stamnes, M.A. and Rothman, J.E. (1993). The binding of AP-1 clathrin adaptor particles to Golgi membranes requires ADP-ribosylation factor, a small GTP-binding protein. *Cell*, **73**, 999–1005.

Stick, R. (1992). The gene structure of Xenopus nuclear lamin A: a model for the evolution of A-type from B-type lamins by exon shuffling. *Chromosoma*, **101**, 566-74.

Stone, E.M., Lotery, A.J., Munier, F.L., Heon, E., Piguet, B., Guymer, R.H., Vandenburgh, K., Cousin, P., Nishimura, D., Swiderski, R.E., Silvestri, G., Mackey, D.A., Hagerman, G.S., Bird, A.C., Sheffield, V.C. and Schorderet, D.F. (1999). A single EFEMP1 mutation associated with both Malattia Leventinese and Doyne honeycomb retinal dystrophy. *Nature Genet.*, **22**, 199-202.

Stone, E.M., Webster, A.R., Vandenburgh, K., Streb, L.M., Hockey, R.R., Lotery A.J. and Sheffield, V.C. (1998). Allelic variation in ABCR associated with Stargardt disease but not age-related macular degeneration. *Nature Genet.*, **20**, 328-329.

Strachan, T. and Read, A.P. (1996). *Human Molecular Genetics*, p.88, BIOS Scientific Publishers Limited, Oxford, UK.

Subbaraya, I., Ruiz, C.C., Helekar, B.S., Zhao, X., Gorczyca, W.A., Pettenati, M.J., Rao, P.N., Palczewski, K. and Baehr, W. (1994). Molecular characterization of human and mouse photoreceptor guanylate cyclase-activating protein (GCAP) and chromosomal localization of the human gene. *J. Biol. Chem.*, **269**, 31080-31089.

Südhof, T. C. (1995) *Nature (London)* 375, 645-653 .Kelly, R. B. (1995) *Curr. Biol.* 5, 257-259.

Sullivan, L.S., Heckenlively, J.R., Bowne, S.J., Zuo, J., Hide, W.A., Gal, A., Denton, M., Inglehearn, C.F., Blanton, S.H. and Daiger, S.P. (1999). Mutations in a novel retina-specific gene cause autosomal dominant retinitis pigmentosa. *Nat. Genet.* **22**, 255-259.

Sung, C.H., Davenport, C.M. and Nathans, J.(1993). Rhodopsin mutations responsible for autosomal dominant retinitis pigmentosa. Clustering of functional classes along the polypeptide chain. *J. Biol. Chem.* **268**, 26645-9.

Sung, C. H., Makino, C., Baylor, D. & Nathans, J. (1994). A rhodopsin gene mutation responsible for autosomal dominant retinitis pigmentosa results in a protein that is defective in localisation to the photoreceptor outer segment. *J. Neurosci.* 14, 5818-5833.

Sun, H. and Nathans, J. (1997). Stargardt's ABCR is localized to the disc membrane of retinal rod outer segments. *Nature Genet.*, **17**, 15-16.

Sundin, O., Yang, J.-M., Li, Y., Zhu, D., Hurd, J.N., Mitchell, T.N., Silva, E.D. and Maumenee, I.H. (2000). Genetic basis of total colourblindness among the Pingelapese islanders. *Nature Genet.*, **25**, 289-292.

Sung, C.-H., Makino, C., Baylor, D. and Nathans, J. (1994). A rhodopsin gene mutation responsible for autosomal dominant retinitis pigmentosa results in a protein that is defective in localization to the photoreceptor outer segment. *J. Neurosci.*, **14**, 5818-5833.

Surguchov, A., Bronson, J.D., Banerjee, P., Knowles, J.A., Ruiz, C., Subbaraya, I., Palczewski, K. and Baehr, W. (1997). The human GCAP1 and GCAP2 genes are arranged in a tail-to-tail array on the short arm of chromosome 6 (p21.1). *Genomics*, **39**, 312-322.

Sutton, R.B., Davletov, B.A., Berghuis, A.,M., Südhof, T.C. and Sprang, S.R. (1995). Structure of the first C_2 domain of synaptotagmin I: a novel Ca^{2+}/phospholipid-binding fold. *Cell*, **80**, 929-938.

Swaroop, A., Xu, J., Pawar, H., Jackson, A., Skolnick, C. and Agarwal, N. (1992). A conserved retina-specific gene encodes a basic motif-leucine zipper domain. *Proc. Nat. Acad. Sci.*, **89**, 266-270.

Szczesny, P.J. (1995). Retinitis pigmentosa and the question of photoreceptor connecting cilium defects. *Graefes Arch. Clin. Exp. Ophthalmol.* **233**, 275-83.

Sorsby, A. (1966). The incidence and causes of blindness in England and Wales 1948-1962. *Reports on Public Health and Medical Subjects* **No. 114**, HMSO, London.

Tai, A.W., Chuang, J.-Z., Bode, C., Wolfrum, U. and Sung, C.-H. (1999). Rhodopsin's carboxy-terminal cytoplasmic tail acts as a membrane receptor for cytoplasmic dynein by binding to the dynein light chain Tctex-1. *Cell*, **97**, 877-87.

Teague, P.W., Aldred, M.A., Jay, M., Dempster, M., Harrison, C., Carothers, A.D., Hardwick, L.J., Evans, H.J., Strain, L., Brock, D.J.H., Bundey, S., Jay, B., Bird, A.C., Bhattacharya, S.S. and Wright, A.F. (1994). Heterogeneity analysis in 40 X-linked retinitis pigmentosa families. *Am. J. Hum. Genet.*, **55**, 105-111.

Thiselton, D.L., Hampson, R.M., Nayudu, M., van Maldergem, L., Wolf, M.L., Saha, B.K., Bhattacharya, S.S. and Hardcastle, A.J. (1996). Mapping the RP2 locus for X-linked retinitis pigmentosa on proximal Xp: a genetically defined 5-cm critical region and exclusion of candidate genes by physical mapping. *Genome Res.*, **6**, 1093-1102.

Thompson J.D., Higgins D.G. and Gibson T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673-4680.

Tonin, P., Shanske, S., Miranda, A.F., Brownell, A.K., Wyse, J.P., Tsujino, S. and DiMauro, S. (1993). Phosphoglycerate kinase deficiency: biochemical and molecular genetic studies in a new myopathic variant (PGK Albert). *Neurol.*, **43**, 387-391.

Tornow J. and Santangelo, G.G. (1990). Efficient expression of *Saccharomyces cerevisiae* glycolytic gene *ADH1* is dependent upon a *cis*-acting regulatory element (UAS_{PRG}) found initially in genes encoding ribosomal proteins. *Gene*, **90**, 79-85.

Trahey, M., Wong, G., Halenbeck, R., Rubinfeld, B., Martin, G.A., Ladner, M., Long, C.M., Crosier, W.J., Watt, K., Koths, K. and McCormick, F. (1988). Molecular cloning of two types of GAP complementary DNA from human placenta. *Science*, **242**, 1697-1700.

Travis, G.H., Brennan, M.B., Danielson, P.E., Kozak, C.A. and Sutcliffe, J.G. (1989). Identification of a photoreceptor-specific mRNA encoded by the gene responsible for retinal degeneration slow (RDS). *Nature*, **338**, 70-73.

Travis, G.H., Christerson, L., Danielson, P.E., Klisak, I., Sparkes, R.S., Hahn, L.B., Dryja, T.P. and Sutcliffe, J.G. (1991). The human retinal degeneration slow (RDS) gene: chromosome assignment and structure of the mRNA. *Genomics*, **10**, 733-739.

Treisman, J.E. (1999). A conserved blueprint for the eye? *BioEssays*, **21**, 843-850.

Tsang, S.H., Gouras, P., Yamashita, C.K., Kjeldbye, H., Fisher, J., Farber, D.B. and Goff, S.P. (1996). Retinal degeneration in mice lacking the gamma subunit of the rod cGMP phosphodiesterase. *Science*, **272**, 1026-1029.

Tucker, C.L., Woodcock, S.C., Kelsell, R.E., Ramamurthy, V., Hunt, D.M. and Hurley, J.B. (1999). Biochemical analysis of a dimerization domain mutation in RetGC-1 associated with dominant cone-rod dystrophy. *Proc. Nat. Acad. Sci.* **96**, 9039-44.

Uberbacher, E.C. and Mural, R.J. (1991). Locating protein-coding regions in human dna-sequences by a multiple sensor neural network approach. *Proc. Nat. Acad. Sci.*, **88**, 11261-11265.

Valle, D. and Simell, O. (1995). The hyperornithemias. *In*: C.R. Scriver, A.L.Beaudet, W.S. Sly, D. Valle, (eds.) *The Metabolic and Molecular Bases of Inherited Disease*, 7th Ed., pp. 1147-1186, McGraw-Hill, New York, USA.

van Aelst, L., Barr, M., Marcus, S., Polverino, A. and Wigler, M. (1993). Complex formation between RAS and RAF and other protein kinases. *Proc. Nat. Acad. Sci.*, **90**, 6213-6217.

van den Born, L.I., Bergen, A.A.B. and Bleeker-Wagemakers, E.M. (1992). A retrospective study of registered retinitis pigmentosa patients in the Netherlands. *Ophthalm. Paediat. Genet.*, **13**, 227-236.

van den Hurk, J.A.J.M., Schwartz, M., van Bokhoven, H., van de Pol, T.J.R., Bogerd, L., Pinckers, A.J.L.G., Bleeker-Wagemakers, E.M., Pawlowitzki, I.H., Ruther, K., Ropers, H.-H. and Cremers, F.P.M. (1997). Molecular basis of choroideremia (CHM): Mutations involving the Rab escort protein-1 (REP-1) gene. *Hum. Mutat.*, **9**, 110-117.

vanDorp, D.B., Wright, A.F., Carothers, A.D. and Bleeker-Wagemakers, E.M. (1992). A family with RP3 type of X-linked retinitis pigmentosa: an association with ciliary abnormailities. *Hum. Genet.*, **88**, 331-334.

van Hooser, J.P., Aleman, T.S., He, Y-G., Cideciyan, A.V., Kuksa, V., Pittler, S.J., Stone, E.M., Jacobson, S.G. and Palczewski, K. (2000). Rapid restoration of visual pigment and function with oral retinoid in a mouse model of childhood blindness. *Proc. Nat. Acad. Sci.*, **97**, 8623-8628.

Verpy, E., Leibovici, M., Zwaenepoel, I., Liu, X.-Z., Gal, A., Salem, N., Mansour, A., Blanchard, S., Kobayashi, I., Keats, B.J.B., Slim, R. and Petit, C. (2000). A defect in harmonin, a PDZ domain-containing protein expressed in the inner ear sensory hair cells, underlies Usher syndrome type 1C. *Nature Genet.*, **26**, 51-55.

Vervoort, R., Lennon, A., Bird, A.C., Tulloch, B., Axton, R., Miano, M.G., Meindl, A., Meitinger, T., Ciccodicola, A. and Wright, A.F. (2000). Mutational hot spot within a novel *RPGR* exon in X-linked retinitis pigmentosa. *Nature Genet.*, **25**, 462-266.

Vidal, M., Brachmann, R.K., Fattaey, A., Harlow, E. and Boeke, J.D. (1996). Reverse two-hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-protein interactions. *Proc. Nat. Acad. Sci.*, **93**, 10315-10320.

Voipio, H., Gripenberg, V., Rraitta, C. and Horsmaanheimo, A. (1964). Retinitis pigmentosa; a preliminary report. *Hereditas*, **52**, 247.

von Heijne, G. (1986) A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.*, **14**, 4683-4690.

von Heijne, G. (1992) Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J. Mol. Biol.*, **225**, 487-494.

Waardenburg, P.J. (1963). *Genetics and Ophthalmology*, p.1579, Charles C. Thomas, Springfield, USA.

Wang, G., Seeidman, M.M. and Glazer, P.M. (1996). Mutagenesis in mammalian cells induced by triple helix formation and transcription-coupled repair. *Science* **271**, 802-805.

Wang, T., Milam, A.H., Steel, G. and Valle, D. (1996). A mouse model of gyrate atrophy of the choroid and retina: early pigment epithelium damage and progressive retinal degeneration. *J. Clin. Invest.*, **97**, 2753-2762.

Wang, T., Steel, G., Milam, A.H. and Valle, D. (2000). Correction of ornithine accumulation prevents retinal degeneration in a mouse model of gyrate atrophy of the choroid and retina. *Proc. Nat. Acad. Sci.*, **97**, 1224-9.

Wang, Y., Okamoto, M., Schmitz, F., Hofmann, K. and Südhof, T. C. (1997). Rim is a putative Rab3 effector in regulating synaptic-vesicle fusion. *Nature*, **388**, 593-598.

Weaver, D.T. and DePamphilis, M.L. (1982). Specific sequences in native DNA that arrest synthesis by DNA polymerase alpha. *J. Biol. Chem.*, **257**, 2075-2086.

Weber, B.H.F., Vogt, G., Pruett, R.C., Stohr, H. and Felbor, U. (1994). Mutations in the tissue inhibitor metalloproteinases-3 (TIMP3) in patients with Sorsby's fundus dystrophy. *Nature Genet.*, **8**, 352-356.

Weil, D., Blanchard, S., Kaplan, J., Guilford, P., Gibson, F., Walsh, J., Mburu, P., Varela, A., Levilliers, J., Weston, M.D., Kelley, P.M., Kimberling, W.J., Wagenaar, M., Levi-Acobas, F., Larget-Piet, D., Munnich, A., Steel, K.P., Brown, S.D.M. and Petit, C. (1995). Defective myosin VIIA gene responsible for Usher syndrome type 1B. *Nature*, **374**, 60-61.

Weitz, C.J., Went, L.N. and Nathans, J. (1992). Human tritanopia associated with a third amino acid substitution in the blue sensitive visual pigment gene. *Am. J. Hum. Genet.*, **51**, 444-446.

Weleber, R.G. and Kennaway, N.G. (1988). Gyrate atrophy of the choroid and retina. *In*: J.R. Heckenlively (ed.) *Retinitis Pigmentosa*, pp.198-220, J.B. Lippincott Company, Philadelphia, USA.

Weleber, R.G., Carr, R.E., Murphey, W.H., Sheffield, V.C. and Stone, E.M. (1993). Phenotypic variation including retinitis pigmentosa, pattern dystrophy, and Fundus Flavimaculatus in a single family with a deletion of codon 153 or 154 of the peripherin/RDS gene. *Arch. Ophthal.*, **111**, 1531-1542.

Weng, J., Mata, N.L., Azarian, S.M., Tzekov, R.T., Birch, D.G. and Travis, G.H. (1999). Insights into the functions of rim protein in photoreceptors and etiology of Stargardt's disease from the phenotype in *ABCR* knockout mice. *Cell*, **98**, 13-23.

Wilkie, S.E., Newbold, R.J., Deery, E., Walker, C.E., Stinton, I., Ramamurthy, V., Hurley, J.B., Bhattacharya, S.S., Warren, M.J. and Hunt, D.M. (2000). Functional characterization of missense mutations at codon 838 in retinal guanylate cyclase correlates with disease severity in patients with autosomal dominant cone-rod dystrophy. *Hum. Mol. Genet.* **9**, 3065-73.

Wilkins, M.R., Gasteiger, E., Bairoch, A., Sanchez, J.-C., Williams, K.L., Appel, R.D., Hochstrasser, D.F. (1998). Protein Identification and Analysis Tools in the ExPASy Server *In*: A.J. Link (ed.) *2-D Proteome Analysis Protocols* (1998). Humana Press, New Jersey, USA.

Wolfrum, U. and Schmitt, A. (2000). Rhodopsin transport in the membrane of the connecting cilium of mammalian photoreceptor cells. *Cell Motil. Cytoskeleton* **46**, 95-107.

Workman, P.L., Lucarelli, P., Agostino, R., Scarabino, R., Sacchi, R., Carapella, E., Palmarino, R. and Bottini, E. (1976). Genetic differentiation among Sardinian villages. *Am. J. Phys. Anthrop.*, **43**, 165-176.

Wright, A.F., Mansfield, D.C., Bruford, E.A., Teague, P.W., Thompson, K.L., Riise, R., Jay, M., Patton, M.A., Jeffery, S., Schinzel, A., Tommerup, N. and Fossarello, M. (1995). Genetic studies in autosomal recessive forms of retinitis pigmentosa. *In*: R.E. Anderson, M.M. LaVail and J.G. Hollyfield (eds.) *Degenerative Diseases of the Retina*, pp.293-302, Plenum Press, New York, USA.

Yamamoto, H., Simon, A., Eriksson, U., Harris, E., Berson, E.L. and Dryja, T.P. (1999). Mutations in the gene encoding 11-*cis* retinol dehydrogenase cause delayed dark adaptation and Fundus Albipunctatus. *Nature Genet.*, **22**, 188-191.

Yamamoto, S., Sippel, K.C., Berson, E.L. and Dryja, T.P. (1997). Defects in the rhodopsin kinase gene in the Oguchi form of stationary night blindness. *Nature Genet.*, **15**, 175-178.

Yan, D., Swain, P.K., Breuer, D., Tucker, R.M., Wu, W., Fujita, R., Rehemtulla, A., Burke, D. and Swaroop, A. (1998). Biochemical characterization and subcellular localisation of the mouse retinitis pigmentosa GTPase regulator (mRPGR). *J. Biol. Chem.*, **271**, 19656-19663.

Yang-Feng, T.L. and Swaroop, A. (1992). Neural retina-specific leucine zipper gene NRL (D14S46E) maps to human chromosome 14q11.1-q11.2. *Genomics*, **14**, 491-492.

Yokota, T., Shiojiri, T., Gotoda, T. and Arai, H. (1996). Retinitis pigmentosa and ataxia caused by a mutation in the gene for the alpha-tocopherol-transfer protein. (Letter) *New Eng. J. Med.*, **335**, 1770-1771.

Yoneda, Y., Hieda, M., Nagoshi, E. and Miyamoto, Y. (1999). Nucleocytoplasmic protein transport and recycling of Ran. *Cell Struct. Funct.*, **24**, 425-433.

Zerial, M., and Stenmark, H. (1993). Rab GTPases in vesicular transport. Curr. Opin. Cell Biol. 5, 613–620.

Zhang, J.Z., Davletov, B.A., Südhof, T.C. and Anderson, R.G.W. (1994). Synaptotagmin I is a high affinity receptor for clathrin AP-2: implications for membrane recycling. *Cell*, **78**, 751-760.

APPENDIX

PUBLISHED MATERIAL

Human Molecular Genetics, 2000, Vol. 9, No. 14 2085–2093

ARTICLE

Identification of a novel protein interacting with RPGR

James P. Boylan and Alan F. Wright*

MRC Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK

Received 4 May 2000; Revised and Accepted 6 July 2000

DDBJ/EMBL/GenBank accession no. AF260257

A novel protein, called RPGRIP, has been identified as interacting with the RPGR protein, which is mutated in a severe form of human retinal degeneration, X-linked retinitis pigmentosa (*RP3* type). The bovine RPGRIP was identified initially by screening for RPGR-interacting proteins with a bovine retina cDNA library using the yeast two-hybrid system. The specificity of the interaction was confirmed by co-immunoprecipitation of *in vitro* translated protein and using RPGR mutants. The human *RPGRIP* gene was isolated and shown to be expressed in retina and testis. Human *RPGRIP* spans a genomic interval of 34 kb, and consists of 15 exons, some of which are alternatively spliced. It was mapped using monochromosomal and radiation hybrid cell lines to chromosomal region 14q11. The function of RPGRIP is unknown; it shows no homology to proteins of known function, although it is predicted to form two coiled-coil domains at the N-terminus. *RPGRIP* is a strong candidate gene for causing human retinal degeneration.

INTRODUCTION

Retinitis pigmentosa (RP) is a heterogeneous group of inherited retinal degenerations which affects 1 in 4000 of the general population (1,2). Over 30 different genes have been implicated in the disease, of which about half have been identified (http://www.sph.uth.tmc.edu/RetNet/disease.htm). X-linked RP (xlRP) affects 10–20% of patients in most populations and is one of the most severe forms of RP, with a prevalence in the region of 1 in 25 000 (3,4). xlRP has been genetically mapped to six loci, *RP2* (5), *RP3* (6), *RP6* (7), *RP15* (8), *RP23* (9) and *RP24* (10). The *RPGR* gene is mutated in the *RP3* form of xlRP, which accounts for 70–80% of patients (11). The N-terminal half of RPGR shows homology to RCC1, a guanine nucleotide exchange factor (GEF) for the small nuclear GTPase Ran, which is concerned with nucleocytoplasmic transport and cell cycle control (6,12–13). *RPGR* shows a complex pattern of alternative splicing, although disease-causing mutations are confined to a single low abundance transcript, consisting of the RCC1-like domain and a novel acidic domain of unknown function (11). The RCC1 domain occurs in several other proteins, including p532 (14), which appears to act as a GEF for Rab and Arf GTPases, suggesting that this domain may generally indicate GEF activity. The crystal structure of RCC1 shows a seven-bladed β-propeller structure, which is similar to the predicted structure of RPGR (15).

The yeast two-hybrid system (16,17) allows the identification of interacting partners for a gene of interest by exploiting the modular nature of transcription factors. Two potentially interacting proteins are expressed in *Saccharomyces cerevisiae*, one of which contains the protein of interest (bait) fused to the DNA-binding domain of a yeast transcription factor such as GAL4, whereas the other is fused to the corresponding activation domain. Binding of the bait to an interacting partner leads to activation of the DNA-binding domain and creates a functional transcription factor, which activates downstream reporter genes.

Recently, the yeast two-hybrid system was used to screen a mouse embryo cDNA library using the first 392 amino acids (RCC1-like domain, RLD) of RPGR as bait (18). The δ-subunit of rod cGMP phosphodiesterase (PDED) was identified as an RPGR-interactor in this way. Rod cGMP phosphodiesterase is a component of the visual transduction cascade of vertebrate photoreceptor cells (19). PDED is an abundant, ubiquitous and highly conserved protein which is proposed to have a role in the solubilization of specific membrane proteins (20). The role of RPGR in this interaction is unknown, but it is interesting that PDED also interacts with small GTPases Rab13 and Arl3 (21).

Here, a bovine retina cDNA library has been screened with the RCC1-like domain of RPGR ($RPGR^{RLD}$) as bait in the yeast two-hybrid system, in order to identify RPGR-interacting proteins. A novel bovine protein was identified which bears no homology to any characterized protein. The human homologue was identified and characterized and its product shown to interact with RPGR in a specific manner, using mutant constructs and *in vitro* co-immunoprecipitation. The human gene is expressed in a subset of tissues, including testis and retina, and was mapped to chromosomal region 14q11.

RESULTS

Yeast two-hybrid library screening

A GAL4-based yeast two-hybrid system was used to search for proteins that interact with RPGR. A bovine retina cDNA

*To whom correspondence should be addressed. Tel: +44 131 467 8437; Fax: +44 131 343 2620; Email: alan.wright@hgu.mrc.ac.uk

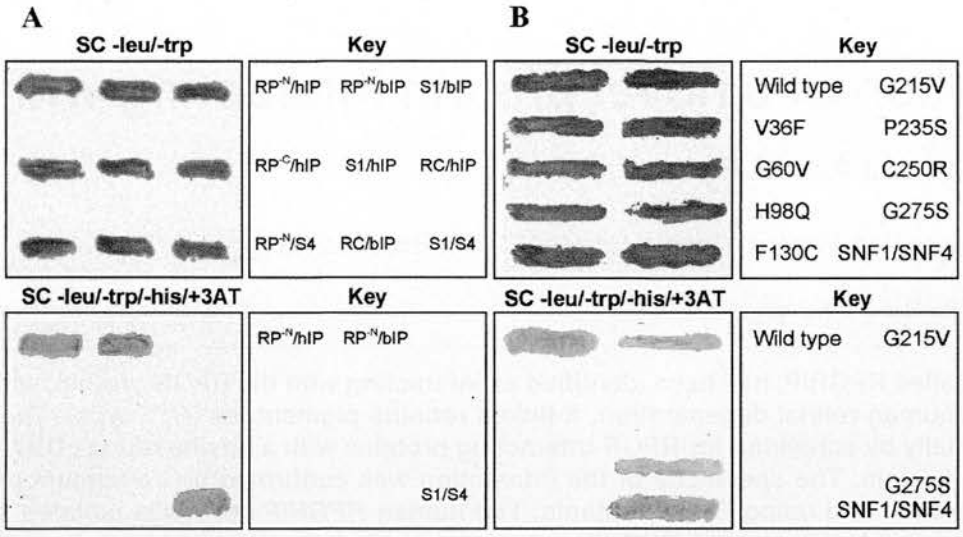


Figure 1. (**A**) A novel protein interacts with the RCC1-like domain of RPGR (RP^{-N}) in the yeast two-hybrid system but does not interact with the C-terminal half of RPGR (RP^{-C}), RCC1 or an unrelated protein (SNF1). Y190 yeast transformed with bait and prey constructs were plated onto SC –leu/–trp to confirm transformation and SC –leu/–trp/–his/+3AT to test for interacting proteins (pAS1 construct/pACTII construct: hIP, human RPGRIP; bIP, bovine RPGRIP; RC, RCC1; S1, SNF1; S4, SNF4). (**B**) The novel human protein interacts with wild-type RPGR and with G215V and G275S RPGR mutants but does not interact with other (V36F, G60V, H98Q, F130C, P235S or C250R) RPGR mutants. Y190 yeast contain the human RPGRIP subcloned into pACTII and pAS1.RPGR mutants. SC –leu/–trp, synthetic complete medium lacking leucine and tryptophan; SC –leu/–trp/–his/+3AT, synthetic complete medium containing 25 mM 3-aminotriazole, lacking leucine, tryptophan and histidine.

library (22) containing 2×10^6 independent clones was screened using the first 12 exons of RPGR, including the whole of the RCC1-like domain ($RPGR^{RLD}$), as bait. The majority of the published RP3 mutations are found in this domain, suggesting an important functional role for this part of the protein (23). Selection for interaction was performed in the Y190 strain, which has conditional *lacZ* and *HIS3* reporter genes. Screening 2×10^6 clones led to the identification of 10 clones showing β-galactosidase activity and histidine prototrophy. Library plasmids were isolated from 8 of the 10 yeast colonies (the remaining 2 were resistant to plasmid recovery) and re-tested for the ability to interact with RPGR. Seven of the eight gave strong positive results and were sequenced. Six of them were found to contain the same insert of 804 bp.

The positive clones were tested for the ability to activate the Y190 reporter genes in the absence of the RPGR bait plasmid and also for the ability to interact with RCC1, Ran and SNF1 (an unrelated protein) (Fig. 1A). The results of these tests were negative, suggesting that the original positive result was specific for and dependent on the presence of RPGR.

Isolation of human *RPGRIP*

A search of the expressed sequence tag (EST) databases for sequences similar to the bovine interactor identified nine human and seven mouse previously uncharacterized homologues, which showed 86% (human) and 76% (mouse) nucleotide identity to the bovine sequence. A human IMAGE EST clone was obtained (accession no. r93221), and the 678 bp insert, which contained an open reading frame (ORF), was subcloned into the pACTII vector. This human clone was tested as a candidate RPGR-interactor in the yeast two-hybrid system (Fig. 1A). It was found to interact with the N-terminal, RCC1-like domain but not with the C-terminal domain of RPGR (the last seven exons), or with RCC1, Ran, PDED or SNF1. In addition, it did not activate the Y190 reporter genes in the absence of a bait construct. The new RPGR-interacting protein was given the name RPGRIP (RPGR-interacting protein).

A partial human RPGRIP (hRPGRIP) cDNA was obtained by PCR amplification of a human testis cDNA using primers derived from overlapping ESTs. 5′ and 3′ rapid amplification of cDNA ends (RACE) experiments were carried out to extend the sequence. Using a primer complementary to the 5′ end of human EST r93221, a 1.2 kb 5′ RACE product was amplified from a human testis Marathon Ready cDNA pool. This was sequenced on both strands and extended the ORF to 1758 bp. The nucleotide context of the first ATG triplet conforms well to the Kozak consensus (24) (Fig. 2). 5′ RACE experiments using human retina cDNA identified a transcript starting 12 bp downstream of the testis transcript. Like the testis 5′ RACE product, the retina transcript contains in-frame stop codons upstream of the translation initiation codon, consistent with it containing the first exon.

Similarly, 3′ RACE experiments identified a transcript extending 84 bp beyond the end of the ORF. Neither of the two most commonly used hexameric polyadenylation signals (AATAAA and ATTAAA) is present at the 3′ end of the cDNA; however, 31 bp upstream of the 3′ end of the RACE product is a hexamer that differs from the canonical AATAAA by a single nucleotide (AGTAAA), which according to Graber *et al.* (25) is the fifth most commonly used hexameric polyadenylation signal. In addition, eight of the nine human ESTs located at

the 3′ end (ai964059, ai655818, ai632512, ai05922, aa782963, aw681763, aa928161 and w26173) terminate 81–83 bp 3′ to the end of the termination codon, consistent with this being the correct polyadenylation signal, since each EST was oligo(dT) primed. The remaining EST (aa476670) is also oligo(dT) primed but ends 315 bp 3′ to the end of the termination codon. Immediately 3′ to the end of this EST is an A-rich region (17/20), suggesting that an extended transcript has false-primed within this region. The length of the cDNA from the start of the 5′-untranslated region (UTR) to the major polyadenylation signal is 1909 bp, of which 1758 bp are an ORF, sufficient to code for a 586 amino acid protein with a predicted size of 67 kDa (Fig. 2).

Genomic structure of *hRPGRIP*

The genomic structure of the *hRPGRIP* gene was obtained from a 196 kb sequence submission (GenBank accession no. AL135744) from a chromosome 14 contig, which includes the entire *hRPGRIP* gene sequence. The exon–intron boundaries and intron sizes are shown in Table 1. Twelve exons were detected initially. The boundaries conform to the GT/AG rule (26), with the exception of the splice donor site at the end of exon 4. The sequence at the exon 4–intron 4 junction was checked repeatedly and confirmed as CTGlgcaagt. It is rare to observe a GC instead of the conserved GT present at most splice donor sites, but other examples of functional splice sites with this sequence have been reported (27).

A panel of bioinformatics programs was applied to the genomic sequence containing *RPGRIP*. The GENSCAN program (28) strongly predicted three extra exons in the middle of intron 4 that had not been detected in cDNAs or RACE products using primers flanking this region. Twelve exons were also predicted to lie upstream of the first detected exon. Primers were designed to anneal to these exons and were used in conjunction with primers from downstream exons. Human retina-derived cDNA was used as the template source in attempts to amplify these exons by PCR. The three exons predicted to lie in intron 4 were detected and maintained the ORF of the shorter transcript, but the 12 upstream exons were not detected. The boundaries of the new exons were sequenced at the genomic level and found to have canonical splice site boundaries (Table 1), as predicted by GENSCAN. The extra exons extend the cDNA from 1946 to 2894 bp and the ORF to 2706 bp, predicting a protein of 902 amino acids and a size of 97 kDa.

The full-length *hRPGRIP* cDNA and intron–exon boundaries are shown in Figure 2A and a schematic representation of the two detected major transcripts in Figure 2B. The arrow in Figure 2B indicates the position of the non-canonical splice site at the exon 3–exon 4 junction. It is interesting to note that the three exons immediately 3′ of this splice site are absent from PCR-amplified *RPGRIP* transcripts.

Interaction of hRPGRIP with mutant RPGR peptides

An experiment was carried out to determine whether the interaction between hRPGRIP and RPGR is disrupted when point mutations associated with xlRP are present in the *RPGR* coding sequence. Eight mutant $RPGR^{RLD}$ constructs were tested in the yeast two-hybrid system. The wild-type $RPGR^{RLD}$ was found to interact with RPGRIP, but the mutants showed either absent (V36F, G60V, H98Q, F130C, G215V, P235S and C250R) or slightly decreased (G215V and G275S) reporter gene activity (Fig. 1B). All of the disease-causing mutations tested, with the exception of F130C, are situated within conserved residues of the $RPGR^{RLD}$ (15), and are therefore likely to be important for correct folding and function.

Co-immunoprecipitation of RPGR and RPGRIP

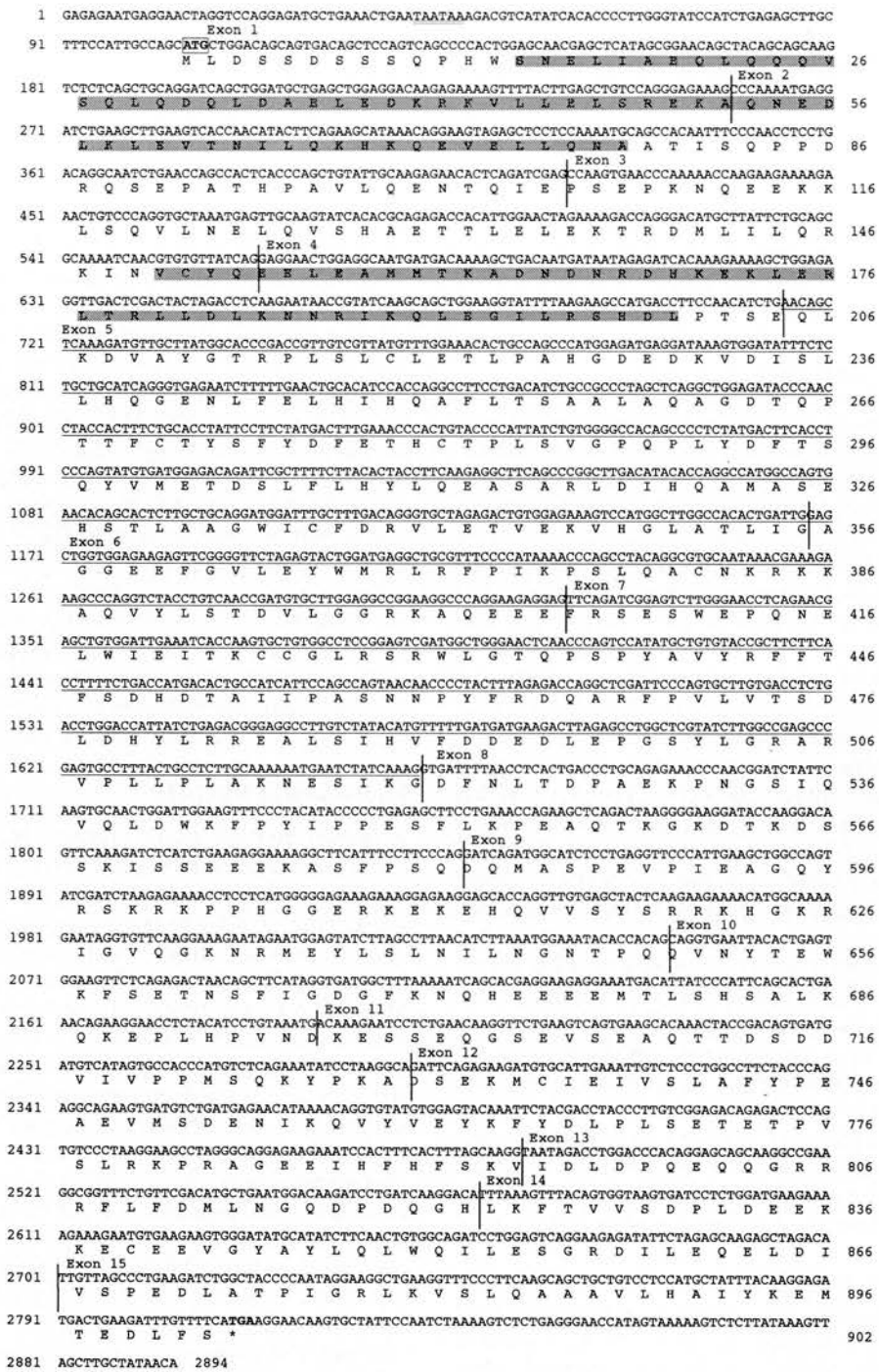
The interaction between RPGR and the RPGRIP was confirmed by *in vitro* co-immunoprecipitation. *In vitro* transcribed/translated and ^{35}S-labelled epitope-tagged RPGR, RPGRIP, Max and lamin proteins were prepared. From a mixture of RPGR-myc and RPGRIP-haemagglutinin (HA), it was possible to co-immunoprecipitate both proteins using either anti-myc or anti-HA antibodies (Fig. 3). From a mixture of RPGR-myc and Max-HA, only RPGR-myc was immunoprecipitated by anti-myc. Similarly, from a mixture of RPGRIP-HA and lamin-myc, only RPGRIP-HA was immunoprecipitated by anti-HA. These results suggest that the interaction between RPGR and RPGRIP is specific and not an artefact of the yeast two-hybrid system. The bands obtained with anti-*c*-myc mouse monoclonal antibody (lanes 1 and 3) are weaker than those with anti-HA rabbit polyclonal antibody (lanes 2 and 4), either because protein G, used to capture antigen–antibody complexes, binds more strongly to rabbit than to mouse antibodies or because the polyclonal antibody binds with higher avidity than the monoclonal (29).

Expression of RPGRIP

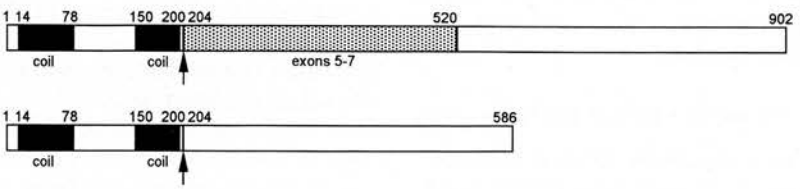
A human multiple tissue northern blot (MTN2) was probed using radiolabelled RPGRIP cDNA. Two different transcripts were detected in the testis lane, a strong band of 2.0 kb and a weak band of 3.1 kb, but no signal was detected in spleen, thymus, prostate, ovary, small intestine, colon or peripheral leukocytes (Fig. 4). A second human multiple tissue northern blot (MTN1) was probed in the same way, and only a very weak 6.2 kb band was identified in the lane containing skeletal muscle RNA, which is presumed to be non-specific, but no signal was obtained in other tissues, including pancreas, kidney, liver, lung, placenta, brain and heart (data not shown). In addition, no signal was obtained with a northern blot containing human retina (data not shown). *RPGRIP* ESTs were identified from a limited range of human tissues, namely fetal liver/spleen (r93221), retina (w28191 and w26173), testis (ai015922 and aa782963), mixed fetal lung, testis and B-cell (aa928161 and aw081763) and pooled germ cell tumour (ai964059, ai655818 and ai632512).

In order to carry out a more sensitive screen for low level *RPGRIP* expression, samples of human and bovine total RNA were used to amplify bovine *RPGRIP* by a single round of RT–PCR. Transcripts were amplified from human retina and testis RNA after one round of amplification (35 cycles) but not from adrenal, brain, heart, kidney, liver, lung or spleen RNAs (Fig. 5A). The lower retina band was sequenced and found to lack exon 12. One round of amplification was sufficient to produce a band of the correct size from the majority of the bovine samples, suggesting that the transcript is more widely expressed (Fig. 5B). When a second round of amplification was carried out (using a single nested primer), a signal was detected in most human tissues (data not shown), as in the case

A


Exon 1
Exon 2
Exon 3
Exon 4
Exon 5
Exon 6
Exon 7
Exon 8
Exon 9
Exon 10
Exon 11
Exon 12
Exon 13
Exon 14
Exon 15

B


1 14 78 150 200 204 520 902
coil
coil
exons 5-7
1 14 78 150 200 204 586
coil
coil

of bovine tissues, but this may be detecting illegitimate transcription products (30).

Chromosomal mapping of *hRPGRIP*

Database searching identified a sequence-tagged site (STS h14a1407, GenBank accession no. G35988) with homology to *RPGRIP*. The STS contains 320 bp of *RPGRIP* genomic sequence covering the intron 12–exon 13 boundary and had been mapped to chromosomal region 14q11 in the course of HAPPY mapping of human chromosome 14 by random fragmentation of haploid cells (31).

A panel of human monochromosomal somatic cell hybrid DNAs was also screened by PCR to determine the chromosome in which *RPGRIP* is located. Human chromosome 14 was found to provide a template for amplification of a genomic fragment of *RPGRIP*, whereas none of the other panel samples, including the rodent control, produced fragments of the expected size.

In order to determine a more precise chromosomal location for *RPGRIP*, a PCR analysis was performed on the HGMP-RC subset of the Genebridge 4 radiation hybrid panel (32). The results were submitted to the Whitehead Institute and showed that *RPGRIP* is located between markers *D14S264* and *D14S275* (6.94 cRays from *D14S264*), which are located in 14q11.

Amino acid sequence analysis

The predicted RPGRIP amino acid sequence was analysed using the HGMP-RC PIX assembly of bioinformatics programs in an attempt to identify possible secondary or tertiary structural features. RPGRIP is predicted to be a soluble protein by the SOSUI system, which analyses proteins on the basis of the physicochemical properties of the amino acid sequence (33). The protein was analysed with the Coils program (34), which predicted RPGRIP to have two coiled-coil domains at the N-terminus, the first encompassing residues 14–78, and the second encompassing residues 150–200 (Fig. 2). BLAST and FASTA searches of dbSPTR identified only a single protein, KIAA1005 (accession no. AB023222), as homologous to RPGRIP. This is an uncharacterized protein of 1055 amino acids, which is 34% identical and 53% similar to RPGRIP (35). No other predicted motifs or domains were identified.

DISCUSSION

The yeast two-hybrid system is a widely used method for detecting interacting proteins and gaining insights into protein function and pathways (36). This is an *in vivo* method for detecting protein–protein interactions and requires confirmation with one or more independent methods, but provides a powerful means of screening large numbers of cDNA clones for interacting products (16,17). The function of the RPGR protein is unknown, although Linari *et al.* (18) previously showed by screening an embryonic mouse cDNA library using the yeast two-hybrid system that it interacts with PDED. This interaction was confirmed by *in vitro* protein binding (pull-down) assays and affinity measurements (K_D = 90 nM), which was consistent with a physiological interaction. Although the function of PDED is still unclear, it has been shown to interact with small GTPases Rab13 and Arl3 (21) and may have a role in the solubilization of Rab13 from membranes (20). It has been postulated that RPGR also interacts with a small GTPase on the basis of its homology to RCC1, a GEF for the Ran GTPase; however, this has yet to be demonstrated.

It is not clear why PDED was not detected on two-hybrid screening of the bovine retina library, since it is known to be expressed in this tissue (37). Possible reasons include differences in the RPGR 'bait', a species difference or the exhaustiveness of library screening. Linari *et al.* (18) used an *RPGR* bait containing amino acids 1–392, whereas here a bait consisting of amino acids 1–502 was used (the RLD consists of residues 39–365). It seems unlikely that this difference would account for the failure to identify PDED. A second possibility is a species difference, since the RPGR–PDED interaction was detected by screening a mouse library with human RPGR, whereas a bovine library was used in this study. Human, mouse and bovine PDE are unusually well conserved, with 98% of amino acids identical between the three species (38). Only one conservative substitution of 150 residues separates bovine from human (Thr68Ala, human:mouse) and three substitutions separate mouse from human (Thr68Ala, Glu10Asp and Arg144Lys). This also seems unlikely, leaving the third possibility, namely a difference in transcript abundance between the mouse embryo and bovine retina libraries. An estimated 2×10^6 bovine retina clones were screened, but it remains a possibility that further screening would identify PDED.

Here, we identify a second interacting protein, RPGRIP, by screening a bovine retinal cDNA library with a clone containing the first 12 exons of human RPGR (RCC1-like domain) using the yeast two-hybrid system. Six of seven strongly positive clones were shown to contain fragments of the bovine *RPGRIP* gene, and the interaction was confirmed with human RPGRIP. Two full-length human cDNAs of 1.95 and 2.89 kb were identified, which appeared to be alternatively spliced orthologues of the bovine gene. The predicted products would yield proteins of 586 and 902 amino acids, respectively, with molecular masses of 67 and 97 kDa.

Further confirmation of the specificity of this interaction came from showing that six of eight xlRP disease-associated RPGR mutations abolish the interaction in yeast two-hybrid experiments. The failure of the G215V and G275S mutants to abolish this interaction, while both abolished interaction with PDED (18), is interesting. Conversely, the V36F variant,

Figure 2. (**A**) cDNA and predicted amino acid sequence of the *RPGRIP* gene (GenBank accession no. AF260257). The translation initiation codon is boxed and an in-frame upstream stop codon is double underlined. The stop codon at the 3′ end of the sequence is in bold. The predicted coiled-coil domains are shaded, and the three alternatively spliced exons (exons 5–7) are underlined. Exon–intron boundaries are shown. (**B**) Schematic representations of the *RPGRIP* transcript with (upper) and without (top) exons 5–7 (bottom). The two predicted coiled-coil-encoding regions are in black and the arrow shows the position of the non-canonical splice site at the exon 4–exon 5 junction. Numbers indicate amino acid positions.

Table 1. Genomic structure of the human *RPGRIP* gene

Exon		Sequence at intron–exon junction		Intron	
Number	Size (bp)	Splice acceptor	Splice donor	Number	Size (bp)
5′-UTR	104	tatcacacccttggggtatccatctgagagcttgctttccattgccag-cATGCTGGGAC			
1	154		AGGGAGAAAGgtaggctggaccttgaagagctctctcaa-atgtggcatcctctacctctg	1	2165
2	161	tggtgataaataactacagaatttcacatttctggattatttttccccagC-CCAAAATGA	TCAGATCGAGgtaagagcctctttaaacaaactagtccact-ctgcagagagatctctggc	2	1080
3	144	tttacctgtcatatttatactccctttaccaatgcgtttccctctacagC-CAAGTGAAC	GTGTTATCAGgtgcaaggaaagatggtacaggaaggggg-atggataacaggaacgtgggaa	3	450
4	151	ttggttttaggccactgagatagaaaagttcagacattattttgtttcag-GAGGAACTGG	CCAACATCTGgcaagtcttagtcctttgttctcctcacttcgg-gaccttccacagctaa	4	2612
5	453	gaaagagctccctacccctttaacggataggcagctttctttcccctcta-gAACAGCTCAAA	ACACTGATTGgtaagtgccgttggcttcctgcggctcctaa-gcaccaatgcagaatttcc	5	161
6	152	cttgccacaccatctgtattccccctgctttcacactttctgctacccag-GAGCTGGTGG	GGAAGAGGAGgtgagaaaaaagatgtgccgaggcatct-cagaggagcctcagccaaacag	6	445
7	343	aacagtctcaagctgccctttttcctcaatccatgaccaacatctttccag-TTCAGATCGG	TCTATCAAAGgtgggagttcgaggttattacatcttcacgc-cctcttcccagtgttgtct	7	1437
8	185	cttgcttatttcatgtgatcaggtcttattaatatctgtttgtttctcagGT-GATTTTAA	TCCTTCCCAGgtaactctccaggactccacaggtagcaga-tctctgccaatcctatggag	8	612
9	204	aggttgttaaactaccagcttgtaatgctatctctgatctttctattcagG-ATCAGATGG	TACACCAGAGgtaagaccttaaaaactctgaagcactaaa-ttgtgggttgtagtgtcccc	9	1618
10	139	gaaaaccaagatattaccagctatgtagtattgttgttcttattctgaag-CAGGTGAATT	CCTGTAAATGgtattgtctttttaaaatctatttttttttcctagca-ctttggggaggccga	10	4215
11	101	ggctaaagtgctttgaaacagttctataactgcaacctcttctctagcag-ACAAAGAATC	TCCTAAGGCAgtaagtacactggagtaatcattgcatacga-gataagacgataaataatg	11	8329
12	193	cttggagcctcactaacctttaggaactaaataaacattttccttatcag-GATTCAGAGA	TTTAGCAAGGgtgaggcatcctgtgtggttactggggtga-ggaagtctgatgaacattga	12	1884
13	85	cactgcaacagtatatgattctttgctttttctcttacccttaatacagTA-ATAGACCT	ATCAAGGACAgtaagcatctgctttccactttgaaacaaag-gagatattgagacagaaat	13	2973
14	131	gaagcattaagagtatcaacagtgctgaattaaatgcaatttctttttag-TTTAAAGTTT	GAGCTAGACAgtgagtcatttttttttcagttctaattatttcca-aggaaatcatcctaa	14	2800
15	113	gtatcaaagtgttcaactgagtgatgctgtttttttccctttcccaacagT-TGTTAGCCC	TTTGTTTTCATGAaagtgctattccaatctaaaagtctctg-agggaaccatagtaaaaggaac		

The exonic sequence is in upper case, and the genomic sequence in lower case. The atypical splice donor site at the exon–intron junction at the end of exon 4 is underlined.

which is associated with X-linked congenital stationary night blindness (T. Meitinger, personal communication), does abolish the interaction with RPGRIP, which it does not for PDED (18). These results suggest that different RPGR subdomains within the RLD are likely to be involved in interaction with RPGRIP and PDED, and that the latter may be more important for full expression of the xlRP disease.

Further evidence relating to the specificity of the RPGR–RPGRIP interaction comes from expression studies. *RPGR* shows multiple alternatively spliced products in human, mouse and bovine tissues, most of which are widely expressed, but one of which (exon ORF14/15 transcript) is expressed exclusively in bovine retina and testis (11). Human retinal RPGR transcripts have only been found using RT–PCR, and show a complex pattern of expression, with two major and several minor transcripts (11). In northern analyses of multiple human tissue RNA samples, *RPGRIP* is only detectable in testis, although it is also readily detected in retina using more sensitive RT–PCR analyses (Fig. 5A). In bovine tissues, RPGRIP expression is detectable by RT–PCR in a wide range of tissues (Fig. 5B). This difference between human and bovine tissues may reflect a genuine species difference or a difference in mRNA quality. Human retina RNA could only be obtained 24 h post-mortem compared with <30 min for bovine tissues. The observation that RPGRIP could be detected in a wide range of human tissues following two rounds of RT–PCR amplification supports the latter possibility, although there is a danger of detecting illegitimate transcription products in this way (30). The predicted full-length RPGRIP cDNAs are 2.89 kb, which includes exons 5–7, and 1.95 kb if these exons are absent, consistent with the observed bands of 2.0 and 3.1 kb detected in testis by northern analysis. Failure to detect a signal in retina by northern analysis may again be related to mRNA quality.

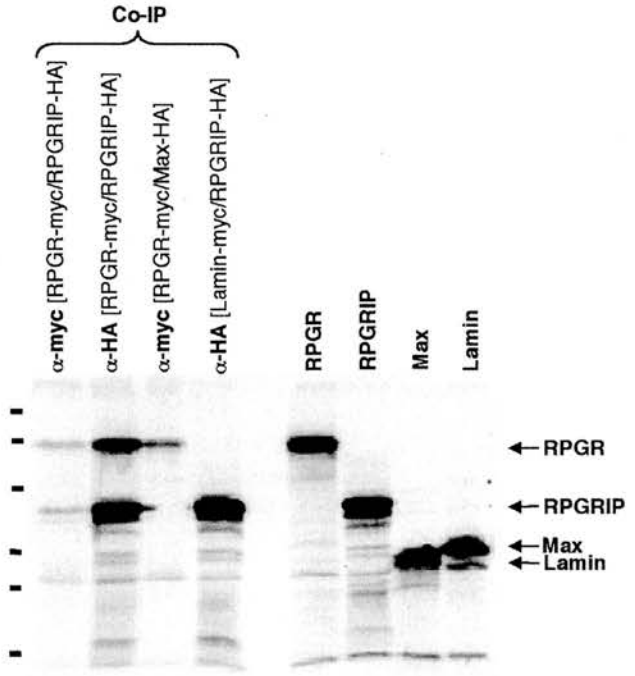


Figure 3. Co-immunoprecipitation of RPGR and RPGRIP *in vitro* transcribed/translated proteins. Lane 1, RPGR-myc plus RPGRIP-HA immunoprecipitated with anti-myc antibody; lane 2, RPGR-myc plus RPGRIP-HA immunoprecipitated with anti-HA antibody; lane 3, RPGR-myc plus Max-HA immunopreciptated with anti-myc antibody; lane 4, RPGRIP-HA plus lamin-myc immunoprecipitated with anti-HA antibody. The four lanes on the right hand side of the gel contain undiluted *in vitro* transcription/translation products for size comparison with proteins on the left hand side. The dashed lines on the left of the gel indicate the position of molecular weight markers; sizes in kDa from top to bottom: 87.0, 51.8, 35.0, 28.4, 20.0 and 7.2.

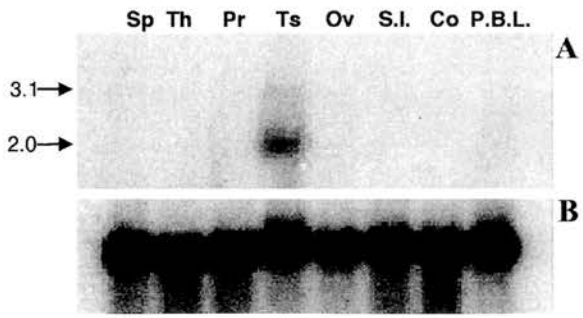


Figure 4. Multiple tissue northern blot probed with (**A**) radiolabelled RPGRIP probe (3′); and (**B**) a glyceraldehyde-3-phosphate dehydrogenase (G3PDH) probe. Sp, spleen; Th, thymus; Pr, prostate; Ts, testis; Ov, ovary; S.I., small intestine; Co, colon; P.B.L., peripheral blood leukocyte. Sizes on the left side of the blot are in kilobases.

The human *RPGRIP* gene is located on the long arm of chromosome 14 in band 14q11, a region containing no known retinal disease genes; however, only a minority of retinitis

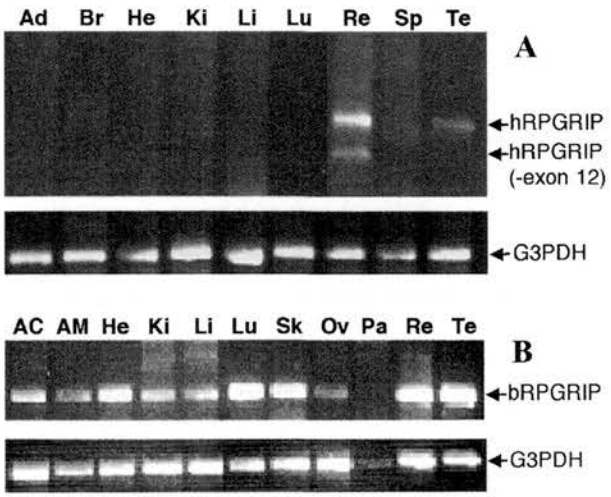


Figure 5. (**A**) Amplification of human RPGRIP (top) and G3PDH (bottom) by RT–PCR from RNA from a variety of tissues (Ad, adrenal; Br, brain; He, heart; Ki, kidney; Li, liver; Lu, lung; Re, retina; Sp, spleen; Te, testis). (**B**) Amplification of bovine RPGRIP (top) and G3PDH (bottom) by RT–PCR from RNA from a variety of tissues (AC, adrenal cortex; AM, adrenal medulla; He, heart; Ki, kidney; Li, liver; Lu, lung; Sk, skeletal muscle; Ov, ovary; Pa, pancreas; Re, retina; Te, testis). The low signal from the pancreatic sample appears to result from poor quality RNA.

pigmentosa patients are accounted for by known or mapped genes (approximately one-third; http://www.sph.uth.tmc.edu/RetNet/disease.htm). The amino acid sequence of RPGRIP gives little indication of its function. Sequence analysis indicates a soluble protein with two coiled-coil domains at the N-terminus of the protein and identifies an uncharacterized human homologue, KIAA1005. Further work will be necessary to elucidate its function, but *RPGRIP* remains a strong candidate gene for human retinal degenerations.

MATERIALS AND METHODS

Two-hybrid screening

A bovine retina cDNA library, kindly provided by Dr C.H. Sung, was used in the GAL4 yeast two-hybrid activation domain vector pACTII as previously described (22). The first 12 exons of *RPGR* were cloned into the two-hybrid bait vector pAS1 (39). The bait construct and the bovine retina library were then co-transfected into *S.cerevisiae* strain Y190 (40). Y190 (Trp^-, Leu^-, His^-, $gal4^-$, $gal80^-$) contains two genomically integrated conditional reporter genes, *HIS3* and *LacZ*, which are under the control of GAL4-responsive upstream activation sites. Transformants were plated onto synthetic complete medium lacking histidine, leucine and tryptophan, and grown at 30°C in the presence of 25 mM 3-aminotriazole. Putative positive colonies were selected by growth on medium lacking histidine and by expression of β-galactosidase activity, as described previously (41). Plasmids were recovered from yeast colonies by glass bead homogenization and phenol–chloroform extraction (42). Library plasmids (containing a gene encoding leucine prototrophy) were amplified selectively in the leucine auxo-

trophic bacterial strain HB101 (Gibco BRL, Paisley, UK) grown on medium lacking leucine. Mutant *RPGR* constructs were obtained from Dr J. Becker (Max Planck Institut für Molekulare Physiologie, Dortmund, Germany) and the inserts subcloned into pAS1. These were then tested in the yeast two-hybrid system for their ability to interact with RPGRIP as described above. Ran and RCC1 were amplified by PCR from a cDNA library, subcloned into pAS1 and sequenced. As the positive control for yeast two-hybrid experiments, Y190 cells were transfected with pAS1.SNF1 and pACTII.SNF4. SNF1 and SNF4 are two yeast proteins known to interact (16).

Co-immunoprecipitation

The inserts from the pAS.RPGR and pACT.RPGRIP constructs were amplified by PCR. One of the primers used in each case contained a non-annealing 5′ end encoding an epitope tag (myc in the case of RPGR and HA for RPGRIP). The amplified sequences were used to produce [^{35}S]methionine-labelled proteins by *in vitro* transcription–translation using a TnT rabbit reticulocyte lysate system (Promega, Southampton, UK). The proteins were mixed and immunoprecipitated by the addition of either anti-myc or anti-HA antibody in the presence of protein G–agarose (Clontech, Basingstoke, UK). Control experiments were performed by mixing RPGR-myc with HA-tagged radiolabelled Max protein and RPGRIP-HA with myc-tagged lamin protein and immunoprecipitating with anti-myc and anti-HA, respectively. The immunoprecipitation products were separated on a 12% polyacrylamide gel, which subsequently was soaked in Amplify Fluorographic Reagent (Amersham Pharmacia Biotech, Little Chalfont, UK). The gel was dried and exposed to Biomax X-ray film (Kodak, Cambridge, UK) overnight.

Northern blot analysis

A PCR fragment encompassing the C-terminal 678 bp of *RPGRIP* was used to prepare a random primed ^{32}P-labelled probe using the High Prime kit (Boehringer Mannheim, Mannheim, Germany). The probe was hybridized to human RNA blots (MTN1 and MTN2; Clontech) using the Ultrahyb (Ambion, Austin, TX) hybridization solution at 42°C overnight. The blot was washed twice for 5 min in 2× SSC, 0.1% SDS at 42°C and then exposed to a phosphor screen overnight.

Reverse transcriptase–PCR

The Access RT–PCR kit (Promega) was used to amplify human and bovine *RPGRIP* from samples of total human and bovine RNA. The reactions were carried out in an MJ Research (Waltham, MA) PTC-200 Peltier thermal cycler using the following programme: 48°C for 45 min, 94°C for 2 min, then 35 cycles of 94°C for 1 min, 53°C (human) or 60°C (bovine) for 1 min and 72°C for 2 min, then 72°C for 10 min.

RACE

Human retina and testis Marathon ready cDNAs (Clontech) were used to obtain 5′ and 3′ RACE products specific for *RPGRIP*. Two rounds of PCR were carried out using nested primers and the Expand High Fidelity proof-reading polymerase mix (Boehringer Mannheim) in a Perkin Elmer-Cetus (Foster City, CA) DNA thermal cycler. The sequences of the RACE products were determined on both strands.

Chromosomal assignment

A human monochromosomal somatic cell hybrid DNA panel (43) [UK Human Genome Mapping Project Resource Centre (HGMP-RC, Cambridge, UK)] and the HGMP subset of the Genebridge 4 radiation hybrid panel (32) were screened by PCR using a Perkin Elmer-Cetus DNA thermal cycler using the following programme: 94°C for 2 min, then 30 cycles of 94°C for 30 min, 52°C for 1 min, 72°C for 1 min, then 72°C for 7 min.

Protein prediction programs

The PIX suite of protein prediction programs was used (HGMP-RC; http://www.hgmp.mrc.ac.uk/Registered/Webapp/pix/) which includes prediction of cell localization [Psort (44)], solubility analysis [SOSUI (33)], BLAST searches against domain databases [SBASE (45) and PRODOM (46)], searches against motif and domain databases [Pfam (47), PRINTS (48), BLOCKS (49) and PROSITE (50)], transmembrane predictions [Tmpred (51), Tmap (52), DAS (53) and Phd (54)] and signal peptide predictions [Sigcleave (55)]. The protein was also analysed with the Coils program, which calculates the probability that a sequence will adopt a coiled-coil conformation (34). The four alternative options were run, i.e. using the MTK scoring matrix weighted and unweighted and using the MTIDK scoring matrix weighted and unweighted.

ACKNOWLEDGEMENTS

We thank C.H. Sung for donating the bovine retina cDNA library, and J. Becker for donating the mutant RPGR constructs. We acknowledge the generous financial support of the Foundation Fighting Blindness, British Retinitis Pigmentosa Society and Guide Dogs for the Blind.

REFERENCES

1. Heckenlively, J.R., Boughman, J.A. and Friedman, L.H. (1988). In Heckenlively, J.R. (ed.), *Retinitis Pigmentosa*. J.B. Lippincott, Philadelphia, PA, pp. 14–24.
2. Van Soest, S., Westerveld, A., De Jong, P.T.V.M., Bleeker-Wagemakers, E.M. and Bergen, A.A.B. (1999) Retinitis pigmentosa: defined from a molecular point of view. *Surv. Ophthalmol.*, **43**, 321–334.
3. Jay, M. (1982) On the heredity of retinitis pigmentosa. *Br. J. Ophthalmol.*, **66**, 405–416.
4. Fishman, G. (1978) Retinitis pigmentosa. Genetic percentages. *Arch. Ophthalmol.*, **96**, 822–826.
5. Schwahn, U., Lenzner, S., Dong, J., Feil, S., Hinzmann, B., van Duijnhoven, G., Kirschner, R., Hemberger, M., Bergen, A., Rosenberg, T. *et al.* (1998) Positional cloning of the gene for X-linked retinitis pigmentosa 2. *Nature Genet.*, **19**, 327–332.
6. Meindl, A., Dry, K., Herrmann, K., Manson, F., Ciccodicola, A., Edgar, A., Carvalho, M.R., Achatz, H., Hellebrand, H., Lennon, A. *et al.* (1996) A gene (*RPGR*) with homology to the RCC1 guanine nucleotide exchange factor is mutated in X-linked retinitis pigmentosa (RP3). *Nature Genet.*, **13**, 35–42.
7. Musarella, M., Anson-Cartwright, L., Leal, S., Gilbert, L., Worton, R., Fishman, G. and Ott, J. (1990) Multiple linkage analysis and heterogeneity testing in 20 X-linked retinitis pigmentosa families. *Genomics*, **8**, 286–296.
8. McGuire, R., Sullivan, L., Blanton, S., Church, M., Heckenlively, J. and Daiger, S. (1995) X-linked dominant cone–rod degeneration: linkage mapping of a new locus for retinitis pigmentosa (RP15) to Xp22.13–22.11. *Am. J. Hum. Genet.*, **57**, 87–94.
9. Hardcastle, A., Thiselton, D., Zito, I., Gorin, M. and Bird, A. (1999a) New genetic loci and gene mutations in X-linked retinal disease. *Invest. Ophthalmol. Vis. Sci.*, **40**, S565.

10. Gieser, L., Fujita, R., Göring, H., Ott, J., Hoffman, D., Cideciyan, A., Birch, D., Jacobson, S. and Swaroop, A. (1998) A novel locus (*RP24*) for X-linked retinitis pigmentosa maps to Xq26–q27. *Am. J. Hum. Genet.*, **63**, 1439–1447.
11. Vervoort, R., Lennon, A., Bird, A.C., Tulloch, B., Axton, R., Miano, M.G., Meindl, A., Meitinger, T., Ciccodicola, A. and Wright, A.F. (2000) Mutational hot spot within a novel RPGR exon in X-linked retinitis pigmentosa. *Nature Genet.*, **25**, 462–466.
12. Ohtsubo, M., Kai, R., Furuno, N., Sekiguchi, T., Sekiguchi, M., Hayashida, H., Kuma, K., Miyata, T., Fukushige, S., Murotsu, T. *et al.* (1987) Isolation and characterisation of the active cDNA of the human cell cycle gene (*RCC1*) involved in the regulation of onset of chromosome condensation. *Genes Dev.*, **1**, 585–593.
13. Sazer, S. and Dasso, M. (2000) The Ran decathlon: multiple roles of Ran. *J. Cell Sci.*, **113**, 1111–1118.
14. Rosa, J.L., Casaroli-Marano, R.P., Buckler, A.J., Vilaro, S. and Barbacid, M. (1996) p619, a giant protein related to the chromosome condensation regulator RCC1, stimulates guanine nucleotide exchange on ARF1 and Rab proteins. *EMBO J.*, **15**, 4262–4273.
15. Renault, L., Nassar, N., Vetter, I., Becker, J., Klebe, C., Roth, M. and Wittinghoffer, A. (1998) The 1.7 Å crystal structure of the regulator of chromosome condensation (RCC1) reveals a seven-bladed propeller. *Nature*, **392**, 97–100.
16. Fields, S. and Song, O. (1989) A novel genetic system to detect protein–protein interactions. *Nature*, **340**, 245–246.
17. Chien, C., Bartel, P., Sternglanz, R. and Fields, S. (1991) The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Natl Acad. Sci. USA*, **88**, 9578–9582.
18. Linari, M., Ueffing, M., Manson, F., Wright, A., Meitinger, T. and Becker, J. (1999) The retinitis pigmentosa GTPase regulator, RPGR, interacts with the δ subunit of rod cyclic GMP phosphodiesterase. *Proc. Natl Acad. Sci. USA*, **96**, 1315–1320.
19. Li, N. and Baehr, W. (1998) Expression and characterization of human PDEδ; and its *Caenorhabditis elegans* ortholog CEδ. *FEBS Lett.*, **440**, 454–457.
20. Marzesco, A.M., Galli, T., Louvard, D. and Zahraoui, A. (1998) The rod cGMP phosphodiesterase δ subunit dissociates the small GTPase Rab13 from membranes. *J. Biol. Chem.*, **273**, 22340–22345.
21. Linari, M., Hazel-Bayer, M. and Becker, J. (1999) The δ subunit of rod specific cyclic GMP phosphodiesterase, PDE δ, interacts with the Arf-like protein Arl3 in a GTP specific manner. *FEBS Lett.*, **458**, 55–59.
22. Kumar, R., Chen, S., Scheurer, D., Wang, Q., Duh, E., Sung, C., Rehemtulla, A., Swaroop, A., Adler, R. and Zack, D. (1996) The bZIP transcription factor Nrl stimulates rhodopsin promoter activity in primary retinal cell cultures. *J. Biol. Chem.*, **271**, 29612–29618.
23. Buraczynska, M., Wu, W., Fujita, R., Buraczynska, K., Phelps, E., Andreasson, S., Bennett, J., Birch, D., Fishman, G., Hoffman, D. *et al.* (1997) Spectrum of mutations in the *RPGR* gene that are identified in 20% of families with X-linked retinitis pigmentosa. *Am. J. Hum. Genet.*, **61**, 1287–1292.
24. Kozak, M. (1987) An analysis of 5′-noncoding sequences from 699 vertebrate messenger-RNAs. *Nucleic Acids Res.*, **15**, 8125–8148.
25. Graber, J., Cantor, C., Mohr, S. and Smith, T. (1999) *In silico* detection of control signals: mRNA 3′-end-processing sequences in diverse species. *Proc. Natl Acad. Sci. USA*, **96**, 14055–14060.
26. Mount, S. (1982) A catalogue of splice junction sequences. *Nucleic Acids Res.*, **10**, 459–472.
27. Senapathy, P., Shapiro, M. and Harris, N. (1990) Splice junctions, branch point sites and exons: sequence statistics, identification and applications to genome projects. *Methods Enzymol.*, **183**, 253–278.
28. Burge, C. and Karlin, S. (1998) Finding the genes in genomic DNA. *Curr. Opin. Struct. Biol.*, **8**, 346–354.
29. Harlow, E. and Lane, D. (1988) *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
30. Chelly, J., Concordet, J.P., Kaplan, J.C. and Kahn A. (1989) Illegitimate transcription: transcription of any gene in any cell type. *Proc. Natl Acad. Sci. USA*, **86**, 2617–2621.
31. Dear, P., Bankier, A. and Piper, M. (1998) A high-resolution metric HAPPY map of human chromosome 14. *Genomics*, **48**, 232–241.
32. Gyapay, G., Schmitt, K., Fizames, C., Jones, H., VegaCzarny, N., Spillett, D., Muselet, D., PrudHomme, J.F., Dib, C., Auffray, C. *et al.* (1996) A radiation hybrid map of the human genome. *Hum. Mol. Genet.*, **5**, 339–346.
33. Hirokawa, T., Boon-Chieng, S. and Mitaku, S. (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics*, **14**, 378–379.
34. Lupas, A., Van Dyke, M. and Stock, J. (1991) Predicting coiled coils from protein sequences. *Science*, **252**, 1162–1164.
35. Nagase, T., Ishikawa, K., Suyama, M., Kikuno, R., Hirosawa, M., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N. and Ohara, O. (1999) Prediction of the coding sequences of unidentified human genes. XIII. The complete sequences of 100 new cDNA clones from brain which code for large proteins *in vitro*. *DNA Res.*, **1**, 63–70.
36. Luban, J. and Goff, S.P. (1995) The yeast two-hybrid system for studying protein–protein interactions. *Curr. Opin. Biotechnol.*, **6**, 59–64.
37. Florio, S.K., Prusti, R.K. and Beavo, J.A. (1996) Solubilization of membrane-bound rod phosophodiesterase by the rod phosphodiesterase recombinant δ subunit. *J. Biol. Chem.*, **271**, 24036–24047.
38. Li, N., Florio, S.K., Pettenati, M.J., Rao, P.N., Beavo, J.A. and Baehr, W. (1998) Characterization of human and mouse rod cGMP phosphodiesterase δ subunit (PDE6D) and chromosomal localization of the human gene. *Genomics*, **49**, 76–82.
39. Harper, J., Adami, G., Wei, N., Keyomarsi, K. and Elledge, S. (1993) The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell*, **75**, 805–816.
40. Flick, J. and Johnston, M. (1990) Two systems of glucose repression of the GAL1 promoter in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **10**, 4757–4769.
41. Schneider, S., Buchert, M. and Hovens, C. (1996) An *in vitro* assay of β-galactosidase from yeast. *Biotechniques*, **20**, 960–962.
42. Hoffman, C. and Winston, F. (1987) A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. *Gene*, **57**, 267–272.
43. Kelsall, D.P., Rooke, L., Warne, D., Bonzyk, M., Cullin, L., Cox, S., West, L., Povey, S. and Spurr, N.K. (1995) Monochromosomal somatic cell hybrids for rapid gene mapping. *Ann. Hum. Genet.*, **59**, 233–241.
44. Nakai, K. and Horton, P. (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem. Sci.*, **24**, 34–35.
45. Murvai, J., Vlahovicek, K., Barta, E., Cataletto, B. and Pongor, S. (2000) The SBASE protein domain library, release 7.0: a collection of annotated protein sequence segments. *Nucleic Acids Res.*, **28**, 260–262.
46. Corpet, F., Gouzy, J. and Kahn, D. (1999) Recent improvements of the ProDom database of protein domain families. *Nucleic Acids Res.*, **27**, 263–267.
47. Bateman, A., Birney, E., Durbin, R., Eddy, S., Howe, K. and Sonnhammer, E. (2000) The Pfam protein families database. *Nucleic Acids Res.*, **28**, 263–266.
48. Attwood, T., Croning, M., Flower, D., Lewis, A., Mabey, E., Scordis, P., Selley, J. and Wright, W. (2000) PRINTS-S: the database formerly known as PRINTS. *Nucleic Acids Res.*, **28**, 225–227.
49. Henikoff, S. and Henikoff, J. (1991) Automated assembly of protein blocks for database searching. *Nucleic Acids Res.*, **19**, 6565–6572.
50. Hofmann, K., Bucher, P., Falquet L. and Bairoch, A. (1999) The PROSITE database, its status in 1999. *Nucleic Acids Res.*, **27**, 215–219.
51. Hofmann, K. and Stoffel, W. (1993) TMbase—a database of membrane spanning protein segments. *Biol. Chem.*, **347**, 166.
52. Persson, B. and Argos, P. (1994) Prediction of transmembrane segments in proteins utilising multiple sequence alignments *J. Mol. Biol.*, **237**, 182–192.
53. von Heijne, G. (1992) Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J. Mol. Biol.*, **225**, 487–494.
54. Rost, B. and Sander, C. (1993) Prediction of protein structure at better than 70% accuracy. *J. Mol. Biol.*, **232**, 584–599.
55. von Heijne, G. (1986) A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.*, **14**, 4683–4690.